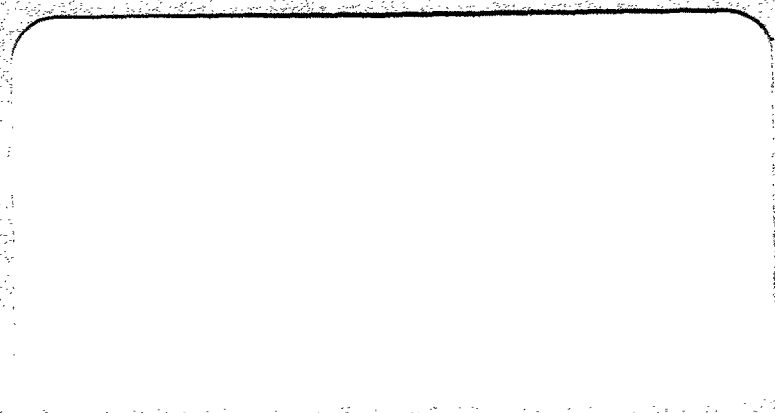
4701



MATHEMATICA
Policy Research, Inc.

Contract No.: 53-3198-O-22
MPR Reference No.: 7925-314

VALIDATION OF THE CHILD-CARE EXPENSE EQUATIONS IN THE MATH MODEL

March 24, 1992

Author:

Nancy Heiser

Submitted to:

U.S. Department of Agriculture
Food and Nutrition Service
3101 Park Center Drive
2nd Floor
Alexandria, VA 22302

Project Officer: Alana Landey

Submitted by:

Mathematica Policy Research, Inc.
600 Maryland Avenue, S.W.
Suite 550
Washington, D.C. 20024

Project Director: Pat Doyle

This work was prepared as one task of a competitively awarded contract; the total amount of the contract is $2,854,698.

ACKNOWLEDGMENTS

The author thanks Robert Cohen, Susan Lund, and Randy Hirscher for conducting computer programming; Pat Doyle and Richard Shin for providing guidance with and reviewing the report; Mindy Weinberg for editing the report; and Sheana Carter for preparing the manuscript.

CONTENTS

CONTENTS (continued)

TABLES

FIGURES

EXECUTIVE SUMMARY

The Food Stamp Program (FSP) provides nutritional assistance nationwide to Fmancially needy households. The FSP is often reviewed and analyzed by policymakers who propose and implement changes to the program. To estimate the budgetary and distributional impacts of proposed and implemented program reforms, the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture has relied primarily on the Micro Analysis of Transfers to Households (MATH)@ model. The MATH model is a microsimulation model that simulates the FSP and the impact of program reforms.

The MATH model is reviewed periodically to ensure its accuracy and validity. This report evaluates the MATH model's procedures for estimating the FSP's dependent-care deduction. Because the database underlying the MATH model, the March Current Population Survey (CPS), does not contain the information on dependent-care expenses that is necessary to calculate the dependent-care deduction, these expenses must be estimated. In this study, we describe the dependent-care deduction modeling procedures in the MATH model, evaluate their performance over time, and discuss the implications of our findings for the MATH model.

The FSP's Dependent-Care Deduction

To be eligible for the FSP, a household must meet asset and income eligibility standards, including a gross and net income standard. Net income is derived from gross income through several allowable deductions. The dependent-care deduction allows households with children and other dependents to deduct expenses of up to a maximum of $160 per dependent per month for caretaking responsibilities, while household members work, look for work, or go to school. The MATH model estimates expenses for children younger than age 15 only, and does not attempt to model expenses for "other dependents", which are not identifiable. Thus, we refer to the deduction and expenses as the "child-care" deduction and "child-care" expenses.

The Child-Care Expense Equations

To estimate child-care expenses, a set of equations were developed (Doyle, Richter, Shin, and Trippe, 1991). The child-care expense equations were estimated for families, since they are the logical decision-making unit for child-care arrangements. The MATH model eventually sums child-care expenses for all families in a household, the unit for which FSP eligibility is determined.

Estimating child-care expenses is a two-stage process. Since all families that have children younger than age 15 and a working parent do not necessarily incur child-care expenses, the model contains a ***probability*** equation to predict whether a family incurs child-care expenses, and an expenditure ***equation*** to predict the amount of child-care expenses for families predicted to incur them.

To estimate the coefficients for these two equations, the 1984 Panel of the Survey of Income and Program Participation (SIPP), Wave 5 was used. SIPP is a longitudinal survey of the civilian, noninstitutionalized population which captures various expenses, monthly income, labor-force participation, and demographic characteristics. The equations consist of characteristics which are

collected in both the CPS and SIPP and which have an impact on or relationship with the probability of a family incurring child-care expenses and the expense amount.

Implementing the Child-Care Expense Equations in the 1991 MATH Model

The equations estimated with SIPP data were incorporated into the 1991 MATH model. This process entailed converting annual income to monthly income, deflating the income data to 1984 for use in the equations (since the child-care expense equations were estimated from 1984 data), inflating the assigned expenses to 1991 dollars (the MATH model simulation year), and summing expenses for all families in a household to obtain household child-care expenses. If the results differed significantly from target dependent-care deduction values obtained from administrative data (Integrated Quality Control Sample data), the results were adjusted or calibrated. This calibration process entailed incorporating additive and multiplicative adjustment factors to the equations and modifying these factors as necessary until the simulated deductions closely met the aggregate targets.

The calibration process for the 1991 MATH model was modified slightly because the predicted 1991 child-care expense controls did not reflect the legislated change in the dependent-care deduction cap from $160 per household to $160 per dependent. Thus, the calibration process first used an alternative model to modify the adjustment factors, using the old dependent-care deduction cap. The targeted dependent-care deduction was $115 over 2.8 percent of all households under the prior legislation. We then used the calibration factors and the new dependent-care deduction cap per dependent. After the calibration process was repeated to reflect this more recent legislation, the final average dependent-care deduction was $122.29 over 2.72 percent of all households.

Evaluation Methodology of the Child-Care Expense Equations

To evaluate the child-care expense equations, we examined how well the equations predicted expenses on the data from which they were estimated (the 1984 Panel of SIPP, Wave 5). We also examined how well the equations performed on more recent SIPP data (the 1987 Panel of SIPP, Wave 6) conducted in late 1988 and early 1989, to examine how well the equations predicted expenses over time. This more recent SIPP data asked the child-care questions of a broader group of families; thus, we also evaluated how well the equations performed on this expanded universe.

We compared various reported outcomes from the SIPP data to predicted outcomes for low-income families, families potentially eligible for the FSP based on their household incomes, and families comprised of at least one person reporting food stamps. For each of these groups of families, we compared (1) the percentage reporting child-care expenses with the percentage predicted to incur child-care expenses, as well as the percentage correctly predicted; (2) average monthly reported and predicted child-care expenses; (3) the distribution of families by the amount of reported and predicted child-care expenses; and (4) the demographic and economic characteristics of families reporting child-care expenses with those predicted to incur child-care expenses. We repeated this analysis using the more recent SIPP data and both the uncalibrated and calibrated equations to determine if the assumptions and relationships contained in the equations are valid based on more recent data and can be used in the next MATH model. We also examined the effect of the random error term contained in the equations.

Evaluation Results of the Child-Care Expense Equations

From our comparison of information reported on the SIPP data from which the equations were estimated with the outcomes predicted by the equations, we found that the equations performed well for low-income families (the universe upon which the equations were estimated), but not as accurately for food stamp families and families potentially eligible for food stamps. Specifically, among low-income families, 27 percent reported child-care expenses averaging $151 per month, and the equations predicted 28 percent to incur child-care expenses averaging $150 per month. Among potentially eligible families, the equations overestimated the percentage of families to incur child-care expenses by 14 percent and average monthly child-care expenses by 5 percent. Among food stamp families, the equations overestimated the percentage of families to incur child-care expenses by 20 percent and the average monthly child-care expenses by 23 percent. Since the equations did not perform as well for the subgroups, the behavior of the subgroups must be different from that of the overall population. To improve the performance of the equations for the potentially eligible subgroup, we could add a dummy variable based on poverty status. However, to control for differences in the behavior of food stamp families, child-care expenses would be dependent on the FSP participation outcome being simulated in the MATH model.

The equations predicted over 70 percent of low-income families to incur or not to incur child-care expenses consistently with the family's reported information. This percentage is a significant improvement over the percentage which would be correctly predicted if the families were simply randomly assigned to incur or not to incur child-care expenses. A comparison of the distributions of families by reported and predicted child-care expenses and of family economic and demographic characteristics for families reporting child-care expenses and those predicted to incur child-care expenses shows that the characteristics are very similar for low-income families, but not as similar for food stamp and potentially eligible families.

Both the child-care probability and expenditure equations contain a random component to control for differences in extraneous characteristics. We analyzed the impact of these random components and found that for low-income families, the random component increased the accuracy of both equations.

Performance of the Child-Care Expense Equations Over Time

By applying the equations to more recent SIPP data (the 1987 Panel of SIPP, Wave 6), we found that the equations did not perform well for any of the groups of families. Specifically, the uncalibrated equation overestimated the percentage of low-income families to incur child-care expenses by 15 percent. This finding suggests that the relationship between the explanatory variables and, the child-care expense status of the family has changed over time. Among food stamp families, the calibrated equation overestimated the percentage of families to incur child-care expenses by 42' percent. Since the calibration factors were developed based on the food stamp families, these results are extremely inaccurate. In fact, the uncalibrated equations performed better than the calibrated equation for food stamp families but not satisfactorily. Among potentially eligible families, neither equation performed well. The uncalibrated equation underestimated the percentage of families to incur child-care expenses by 11 percent, and the calibrated equation overestimated the percentage of families to incur child-care expenses by 29 percent.

In summary, for low-income families, the uncahbrated equation did not perform well and the calibrated equation performed even less well. For food stamp families, we would expect the

calibrated equation to perform well, but it did not, and the uncalibrated equation performed better, but not satisfactorily. Among potentially eligible families, the uncalibrated equation performed better than the calibrated equation, but still not satisfactorily.

On a more positive note, the equations predicted about 70 percent of the cases correctly, which compares favorably to the findings based on the 1984 SIPP data. Similarly, the distributions of the families by the amount of reported and predicted child-care expenses were similar for low-income families, but not as accurate for food stamp or potentially eligible families. For low-income families, the demographic and economic characteristics of families reporting child-care expenses closely reflect those of families predicted to incur child-care expenses. For food stamp families and potentially eligible families, the characteristics of families reporting child-care expenses differ in some ways from those predicted to incur child-care expenses. Specifically, the equations underestimated the percentage of families with single or single-female parents, and overestimated the percentage of families with a nonwhite parent to incur child-care expenses.

Since the equations did not perform well on more recent data for the overall population or the subgroups, the behavior of families reporting child-care expenses most likely changed over time. A comparison of the characteristics of low-income families reporting child-care expenses in the earlier SIPP data to those reporting child-care expenses in the more recent SIPP data revealed that the latter group had less income and were less likely to be employed, but had higher -child-care expenses. Based on their income and employment status, the equation developed on the earlier data would predict lower child-care expenses, and hence the equations underestimated child-care expenses on the more recent SIPP data.

Implications of the Evaluation of the Child-Care Expense Equations for the MATH Model

The child-care expense equations did not perform well on the 1987 SIPP Panel, Wave 6. This finding suggests that the relationship between the explanatory variables in the child-care expense equations and the dependent variables has changed over time. Therefore, we recommend using the existing equations and reestimating the coefficients so that the equations represent more precisely the relationship between the explanatory and dependent variables.

This reestimation would take into account the changes in behavior of low-income families over time, as well as any impact the change in SIPP's child-care questions may have had on the performance of the equations. Although we cannot be any more confident that updating to the 1987 SIPP Panel, Wave 6 will predict child-care expenses in a future year better than the existing equations, we would not recommend abandoning these estimation procedures based on the outcome of one application only, especially since the quality of the data may have changed between the surveys, and because better alternatives do not exist. We would also add a dummy variable (and perhaps some interaction variables) indicating potentially eligible families which would improve the equations' performance for this group.

In addition, there is an underlying discrepancy between the IQCS and the SIPP data, with about one-third of the families in SIPP reporting child-care expenses, compared with only three percent in the IQCS data. Thus, we recommend investigating the discrepancy between the SIPP and IQCS data sets to determine if the relationship between the explanatory and dependent variables differ for the two data sets. Because the IQCS data do not contain sufficient information for the equations to operate, we recommend a simpler comparison of the relationship of various family characteristics to their child-care expenses in the IQCS and in SIPP.

I, INTRODUCTION

The Food Stamp Program (FSP) provides nutritional assistance nationwide to financially needy households. In summer 1990, the FSP served an average of 20 million persons per month at an annual cost of over $13 billion. Since the FSP is the largest and most universally available assistance program nationwide, it is closely scrutinized by policymakers and analysts, who propose and implement changes to the program. To estimate the budgetary and distributional impacts of proposed and implemented changes to the FSP, the U.S. Department of Agriculture's Food and Nutrition Service (FNS) has maintained and enhanced the Micro Analysis of Transfers to Households (MATH)@ model. The MATH model is a microsimulation system that operates with data from the March Current Population Survey (CPS). The CPS database provides information on the annual income, labor-force participation, and demographic characteristics of a large sample of the civilian, noninstitutionalized population and is the best available source of data for simulating the impacts of program reforms.

The MATH model is reviewed periodically to ensure its validity and accuracy. This study continues efforts to validate the MATH model by evaluating the procedures used to estimate the FSP's dependent-care deduction. These procedures compensate for the lack of information on deductible dependent-care expenses in the CPS. This study describes the development and dependent-care deduction modeling procedures used in the MATH model, evaluates how well they predict actual dependent-care expenses and their performance over time, and discusses the implications of our findings for the MATH model. In the remainder of this introductory chapter, we provide an overview of (1) the simulation of the FSP within the MATH model, (2) the procedures used in the MATH model to estimate the dependent-care deduction, and (3) the report's contents.

A. SIMULATION OF THE FOOD STAMP PROGRAM IN THE MATH MODEL

The MATH model estimates the impact of proposed or implemented reforms on the FSP. To support estimating the impact of reforms on the FSP, the data underlying the MATH model are enhanced and the FSP is simulated under legislation for a given time period Using this enhanced data and the simulated **FSP** as a baseline, the model simulates the impact of reforms on the **FSP.** Below, we discuss the data enhancements and the **simulation** process.

1. Enhancements to the March CPS 1988 Data

A current, nationally representative survey that captures detailed demographic and economic characteristics of the population is the primary requirement for developing the baseline data to support simulating program changes. None of the currently available survey databases provides all of the information necessary to represent current and potential FSP households accurately. Although the CPS has major limitations for purposes of simulating eligibility and participation in the FSP, it is a large sample and the best available source of much of the necessary demographic and economic information for households, families, and individuals. Thus, the CPS is the database underlying the MATH model. The input file used to create the latest MATH database is the March supplement to the 1988 CPS.

In its raw form, the CPS does not meet all of the necessary requirements for the MATH database; thus, it must be edited, augmented, and projected to the simulation analysis year (1991). The major steps in constructing the 1991 MATH model's database are as follows:

- Converting the March 1988 CPS file to MATH format
- Creating basic definer variables, identifying persons to be included in non-public assistance filing units, and allocating lump-sum income types to detailed sources
- Projecting the data to April 1991 to capture the impact of projected changes in demographic and economic indicators
- Simulating federal income tax liability and payroll taxes

- Estimating assets
- Simulating eligibility, potential benefits, and participation in public assistance programs--namely, the Aid to Families with Dependent Children (AFDC), Supplemental Security Income (SSI), and General Assistance (GA) programs--to compensate for the underreporting of cash welfare
- Simulating eligibility, potential benefits, and participation in the FSP

Simulating the FSP eligibility, benefits, and participation is a four-step process, involving further data enhancements. First, since income eligibility for food stamps is determined on a monthly basis, the annual income data provided in the CPS must be converted into monthly income amounts. This process is explained and evaluated in Doyle and Trippe (1991). Second, the model classifies persons into food stamp units. Households comprised entirely of AFDC or SSI participants are automatically eligible for food stamp benefits without regard to food stamp income and asset requirements.

Third, the model simulates eligibility and potential benefits for the FSP. Because the CPS does not contain all of the information necessary to simulate eligibility and potential benefits for the FSP, the model enhances the data. The model incorporates equations to compensate for the lack of information in the CPS on the dependent-care, shelter, and medical expenses. *This process of enhancing the data includes the incorporation of the child-cm equations discussed in this report. The* model also estimates countable assets to simulate eligibility for the program. The process of simulating the FSP entails computing monthly gross and net income based on the monthly income amounts and in accordance with program regulations enacted for 1991. Eligibility for the FSP is based on gross and net income conditional on household size, and on assets accessible to the household Net income is derived from gross income having several allowable deductions. Food stamp benefits to eligible households are equal to the maximum benefit amount (set by Congress) minus 30 percent of net income. FNS projected the income limits and maximum benefit amounts used in the 1991 model.

The fourth and final step in simulating the FSP entails selecting eligible households to participate in the program. This final step is explained and evaluated in Martini (1991).

The final product of the MATH system is a synthetic file that represents the demographic and socioeconomic characteristics of a large sample of individual families and associated persons during a given month (April 1991). Using the April 1991 database as a baseline, the model can simulate future program reforms and their aggregate and distributional impact on the program and population. This process is described briefly below.

2. Simulating Reforms to the FSP

To estimate the impact of proposed or implemented reforms on the FSP the MATH model uses a two-step procedure. The model first applies FSP eligibility criteria to households in the CPS, as if they had entered the welfare office to apply for food stamp benefits. Since not all eligible households choose to participate, the MATH model invokes a behavioral model to simulate participation among the households that are considered to be eligiile. Second, the model simulates proposed changes to the FSP by incorporating these proposed policy reforms and repeating the eligibility and participation simulation (using a reform participation algorithm). The simulated current law and proposed program are compared at the household level to determine which households would gain and which would lose benefits under the reform, and at the macro level to determine the net impact of the reform on overall program costs and caseload.

B. ESTIMATING THE DEPENDENT-CARE DEDUCTION IN THE MATH MODEL

This report focuses on how the dependent-care deduction is estimated in the MATH model. In this section, we provide an overview of the FSP's dependent-care deduction and the modeling procedures used in the MATH model to estimate the deduction.

1. The FSP's Dependent-Care Deduction

To be eligible for the FSP, a household must meet asset, gross, and net income eligibility standards. The FSP permits several deductions from household gross monthly income to arrive at the net monthly income. One of these deductions is the dependent-care deduction. Households with dependents receive a dependent-care deduction for expenses incurred in caring for children and other dependents while household members work, seek employment, or attend school. The dependent-care deduction is equal to a household's dependent-care expenses up to a maximum. The maximum dependent-care deduction was $160 *per* household per month until Congress expanded the maximum in the Hunger Prevention Act of 1988 to $160 *per dependent* per month. (For more information on eligibility standards and benefit computation, see Heiser, 1991.)

2. Child-Care Expense Equations in the MATH Model

The MATH model incorporates procedures which impute dependent-care expenses. Based on these estimated expenses, it calculates the dependent-care deduction. The MATH model estimates expenses only for children younger than age 15, and does not estimate expenses for "other dependents" older than age 15 which are not identifiable on the CPS. From this point on, we will refer to the deduction and expenses as the "child-care" deduction and "child-care" expenses.

Imputing child-care expenses is a two-stage process. First, since all families that have children and a working parent do not necessarily incur child-care expenses (because child care can be provided at no cost by friends or relatives), the model predicts whether a given family incurs child-care expenses, and, for those families predicted to incur child-care expenses, the model estimates the amount of the expenses. Thus, this modeling procedure contains one equation for each stage: (1) the *probability* equation predicts whether a family incurs child-care expenses, and (2) the *expenditure* equation predicts the monthly amount of the child-care expenses for families predicted to incur them. The outcome of each of these equations depends on the family's demographic and economic

characteristics. The MATH model equations used to predict deductible child-care expenses are discussed in more detail in Appendix A.

3. Estimating and Implementing the Child-Care Expense Equations in the MATH Model

Since the CPS does not contain information on deductible child-care expenses, another database which contains such information, as well as data on demographic and economic family characteristics available in the CPS, must be used to estimate the coefficients for the two equations. Two complementary sources of information on child-care expenses are the Survey of Income and Program Participation (SIPP) and the Integrated Quality Control System (IQCS) sample. The IQCS offers administrative data which provide information on the deductible child-care expenses of food stamp households. SIPP measures child-care expenses for all households in the United States..

The IQCS data are more precise than the SIPP data for current food stamp households but because the IQCS data do not contain information on the eligible nonparticipating population, these data do not support full simulation of the impact of reforms on child-care deductions. Conversely, although SIPP may provide slightly less precise information on deductible child-care expenses for food stamp households, it describes deductible child-care expenses for all low-income households rather than just those currently participating in the FSP. Therefore, SIPP was used to obtain an estimate of the relationship between the demographic and economic characteristics of a family and its child-care expenses. The IQCS data were eventually used to ensure that the child-care deductions estimated in the MATH model match the observed deductions for participating FSP households.

SIPP is a nationally representative longitudinal survey of the civilian, noninstitutional&d population in the United States. The survey has been administered by the U.S. Bureau of the Census since 1983; replacement panels (or samples) of respondents are added each year. In each round (or "wave") of interviewing, a core set of questions are used to collect information on each of four months prior to the interview date. These core questions yield detailed information on monthly program participation, demographic characteristics, household composition, and income. In most

waves, the monthly core questions are supplemented with questions on topical issues that vary from interview to interview. Because the interviews are staggered, the reference period covered in a given wave is not the same for all sample members. However, all waves contain one calendar month common to the reference period of all sample members which provides a common reference point.

In the 1984 SIPP Panel, each household in the sample was scheduled to be interviewed at four-month intervals over a period of two-and-a-half years beginning in October 1983. The reference periods for the questions in Wave 5 are four-month periods from September 1984 to March 1985. SIPP periodically supplements the core survey by surveying a sample of households about specific topics. In the 1984 SIPP Panel, Wave 5, one such supplement, or topical module, asked questions about the child-care expenses working parents of children younger than age 15.[1]

The child-care equations used in the MATH model were developed based on the data from the 1984 SIPP Panel, Wave S (Doyle, Richter, Shin, and Trippe, 1991). The universe of analysis for this process consisted of families with working parents of children younger than age 15. The family rather than the household was chosen as the unit of analysis because the family is the more likely decision making unit for child-care arrangements and expenses. In addition, since the MATH food stamp model screens out high-income households, we screened out high-income households based on their *monthly* income. Based on this universe, we estimated child-care expense equations which reflect the relationship between the characteristics of a family and its child-care expenses.

After developing the equations, they were implemented in the 1991 MATH model to estimate child-care expenses and the FSP child-care deduction. This implementation process entailed (1) summing predicted child-care expenses for all families in the household to the household level in order to compute the child-care deduction and (2) inflating child-care expenses from 1984 dollars to 1991 dollars. Using the predicted 1991 child-care expenses for food stamp households, the model computes a child-care deduction. The average deduction and the percentage of households entitled

[1]The parent refers to the parent or guardian If two parents are present, the mother is assigned as the parent.

to the deduction computed in the 1991 MATH model deviated somewhat from the targeted values reflected in the FSP administrative (IQCS) data. Therefore, the child-care equations were adjusted, or calibrated, to meet more closely these targeted amounts based on the IQCS data. The calibration process entailed implementing additive and multiplicative factors to the equations and modifying these factors until the simulated deduction closely met the targets based on IQCS data.

C. OVERVIEW OF THE REPORT

In this report, we describe the development of the child-care expense equations from SIPP data (Chapter II) and evaluate how well the equations predict family child-care expenses compared with information reported on the SIPP file (Chapter III). Then, we describe the implementation of these equations in the MATH model and the calibration process undertaken to adjust the child-care deduction information to meet the targets obtained from the IQCS data (Chapter IV). We then evaluate how well the equations (with and without the calibration factors) perform with more recent SIPP data, the 1987 SIPP Panel, Wave 6 conducted in late 1988 and early 1989, to determine whether the equations are accurate based on more recent data (Chapter V). We then discuss the implications of our findings for the child-care expense equations used in the MATH model (Chapter VI).

II. DEVELOPMENT OF TH-E CHILD-CARE EXPENSE EQUATIONS WITH THE 1984 SIPP PANEL, WAVE 5

The ultimate objective of the MATH model is to simulate food stamp eligibility, potential benefits, participation, and the distributional impact of program reforms accurately. Since the data underlying the MATH model, the CPS, do not contain all of the information necessary for simulating eligibility and potential benefits for the FSP, the- model enhances the data to estimate the missing information. To estimate the deductible child-care expenses necessary for calculating the dependent-care deduction in the FSP, SIPP data from the 1984 Panel, Wave 5 was used. The development of the child-care expense equations is documented in Doyle, Richter, Shin, and Trippe (1991) and is reviewed here for background purposes. The child-care expense equations represent the relationship between both the probability of a family's incurring child-care expenses and the amount of monthly child-care expenses, and the demographic and economic characteristics of a family. In this chapter, we describe the SIPP analysis file and the estimated child-care expense equations.

A. THE SIPP ANALYSIS FILE

In the 1984 SIPP Panel, Wave $3,330 unweighted families had a working parent of children younger than age 15. The sample size was reduced to 1,766 families (or 9.3 million families weighted) after high-income families were screened out for consistency with the income screen in the MATH model. Among these families, 477 (2.5 million weighted), or 27 percent, reported child-care expenses that averaged $34.74 per week, or $151 per month.

Based on these data, a probability and an expenditure equation were estimated. The final variables used in both the probability and expenditure equations are described in Table II.1. The estimated coefficients of each equation are presented in Tables IL2 and II.3. (Appendix B provides instructions for interpreting Tables IL2 and II.3). The variables are restricted to those available on both the CPS and SIPP. Each of the equations is described in detail below.

TABLE II.1

VARIABLES USED IN THE CHILD-CARE PROBABILITY AND EXPENDITURE EQUATIONS

Variable	Description
AAC5T7	Average age of children in the family ages 5 to 7. Continuous value.
AAGECLD4	Average age of children in the family ages 0 to 4. Special value. 0 = no children in the age range.
AGE!4	Parent's age. Special value. 85 = 85 years old or older.
AGEGT44	Parent's age is more than 44. 1 = Parent's age is more than 44. 0 = Otherwise.
AGELT25	Parent's age is less than 25. 1 = Parent's age is less than 25. 0 = Otherwise.
AGE35T44	Parent's age is 35 to 44. 1 = Parent's age is 35 to 44. 0 = Otherwise.
EARNPHR1	Parent's total earnings divided by total number of hours worked.
EMPSTAT	All parents are employed. Always equal to **1** for single parents. 1 = Yes. 0 = No.
FEMALE	Parent's gender 1 = Parent is female. 0 = Parent is male.
FEMSING	Parent's gender and marital status 1 = Parent is female and single. 0 = Parent is not female and single.
FNUMPER4	Number of persons in family.
FULLPT	Parent works 35 hours per week or more. 1 = Parent works 35 hours per week or more. 0 = Parent works less than 35 hours per week.

TABLE II.1 (continued)

Variable	Description
HIGRADE	Highest grade attended by parent. 0 = Not applicable if under 15, did not attend, or attended kindergarten only. 1-8 = Elementary. 9-12 = High school. 13-18 = College.
HNF4	Number of families and pseudo families in this household
HRSWRK	The usual number of hours worked per week by the parent.
IRENT	Family rents housing. 1 = Does rent. 0 = Does not rent.
LNBARNPH	Natural logarithm of the parent's total earnings divided by the total number of hours worked.
LNTEARN	Natural logarithm of total earned income of the parent and spouse.
LNTUNERN	Natural logarithm of the total unearned income of the parent and spouse.
NCHLD4	Number of children in the family ages 0 to 4. 0 = no children in the age range.
NCH5T7	Number of children in the family ages 5 to 7. 0 = no children in the age range.
NCH5T11	Number of children in the family ages 5 to 11. 0 = no children in the age range.
NCH12T14	Number of children in the family ages 12 to 14. 0 = no children in the age range.
NCH12T18	Number of children in the family ages 12 to 18. 0 = no children in the age range.
NCH15T18	Number of children in the family ages 15 to 18. 0 = no children in the age range.
NONWHITE	Parent's Race 1 = Parent is not white. 0 = Parent is white.
NUMSFWCH	Number of families eligible for child care in the household.
POSTHS	Parent attended some college. 1 = Parent attended some college. 0 = Otherwise.
SEARNPHR	Spouse's total earnings divided by total number of hours worked.

TABLE II.1 (continued)

Variable	Description
SELEMPL	The parent is self-employed. 1 = Self-employed only. 0 = Otherwise.
SINGLE1	Parent's marital status 1 = Parent is single. 0 = Spouse is present.
SOMEHS	Parent attended some high school. 1 = Parent attended some high school. 0 = Otherwise.
SPOSTHS	Spouse attended some college. 1 = Spouse attended some college. 0 = Otherwise.
SSOMEHS	Spouse attended some high school. 1 = Spouse attended some high school. 0 = Otherwise

SOURCE: Doyle, Richter, Shin, and Trippe (1991).

TABLE IL2

THE RESULTS OF THE CHILD-CARE PROBABILITY EQUATION

Variable	Coefficient	Standard Error	T-ratio	Signif. Level	Mean of x	Std. Dev. of x
Intercept Term						
ONE	-2.31871	.651409	-3.560	.00037	1.0000	.00000
Family Characteristics						
FNUMPER4	-.114712	.418458E-01	-2.741	.00612	4.1421	1.5373
HNF4	-.427506	.129213	-3.309	.00094	1.1636	.44257
N-UMSFWCH	.470165	.293671	1.601	.10938	1.0193	.14931
NCHLD4	.584054	.853592E-01	6.842	.00000	.51812	.68033
NCH5T7	.172842	.153676	1.125	.26071	.36976	.58185
NCH12T14	-.251166	.979818E-01	-2.563	.01037	.42412	.62404
NCH15T18	-.301502	.115194	-2.617	.00886	.27010	.56537
AAGECLD4	.164798	.335475E-01	4.912	.0000O	.92818	1.3776
AAC5T7	.503171E-01	.319334E-01	1.576	.11510	1.9074	2.8209
Demographic Characteristics						
FEMALE	.769883	222022	3.468	.00053	.94168	23442
SINGLE1	.271770	.118126	2.301	.02141	.36070	.48034
NONWHITE	.135298	.962616E-01	1.406	.15987	22763	.41942
AGE4	-.477598E-01	.116324E-01	-4.106	.00004	32.899	7.54%
AGEGT44	.473823	320840	1.477	.13972	.73046E-01	.26029
AGE35T44	.264567	.145992	1.812	.06996	.31144	.46321
SOMEHS	-.607807	283347	-2145	.03195	.64496	.47866
POSTHS	-.891043	.359162	-2.481	.01311	.31087	.46298
HIGRADE	.117231	.335742E-01	3.492	.00048	12.296	2.2435
HRSWRK	.860269E-02	.607500E-02	1.416	.15675	33.438	12.242
SELEMPL	-.811150	.195128	-4.157	.00003	.88901E-01	.28468
FULLPT	.320918	.149261	2.150	.03155	.63194	.48242
Economic Characteristics						
EMPSTAT	392261	.166093	2.362	.01819	.91563	.27802
LNTEARN	.474803E-01	.487074E-01	.975	.32966	6.8854	1.2261
LNTUNERN	.137231E-02	.169746E-01	.081	.93557	2.2455	2.4954
LNEARNPH	.152373	.839055E-01	1.816	.06937	1.4481	.62475
IRENT	-.129569	.830320E-01	-1.560	.11865	.39298	.48855

SOURCE: Doyle, Richter, Shin, and Trippe (1991).

TABLE II.3

THE RESULTS OF THE CHILD-CARE EXPENDITURE EQUATION

Variable	DF	Coefficient	Standard Error	T for HO: Parameter =0	Signif. Level
Intercept Term					
INTERCEP	1	2.048683	0.24362207	8.409	0.0001
Family Characteristics					
NCHLD4	1	0.209042	0.04749948	4.401	0.0001
NCH5T11	1	-0.024290	0.03825647	-0.635	0.5258
NCH12T18	1	-0.135856	0.05162501	-2.632	0.0088
Demographic Characteristics					
FEMSING	1	-0.310098	0.13721251	-2.260	0.0243
NONWHITE	1	-0.140724	0.06166225	-2.282	0.0229
AGELT25	1	-0.088556	0.07060498	-1.254	0.2104
SSOMEHS	1	-0.401719	0.13513529	-2.973	0.0031
SPOSTHS	1	-0.289455	0.14213983	-2.036	0.0423
HIGRADE	1	0.016885	0.01304659	1.294	0.1962
HRSWRK	1	0.025848	0.00264766	9.763	0.0001
Economic Characteristics					
LNTUNERN	1	0.020870	0.01128678	1.849	0.0651
EARNPHR1	1	0.049119	0.01074706	4.570	0.0001
SEARNPHR	1	0.022225	0.01025486	2.167	0.0307
IRENT	1	0.107998	0.05473629	1.973	0.0491

SOURCE: Doyle, Richter, Shin, and Trippe (1991).

NOTE: Adjusted R-square is 0.2597.

B. THE CHILD-CARE PROBABILITY EQUATION

The probability that a qualifying family (one with a working parent and children) incurs positive child-care expenses is a function of the various demographic and economic characteristics of the family. The estimated probability equation approximates this function. To specify a reasonable set of explanatory variables for this equation, Wald or likelihood ratio statistics, theoretical reasoning, and prediction results were used.

Based on these demographic and economic characteristics and their associated coefficients which represent their relationship to a family's child-care expense status, the probability equation computes an index value. This index value is then compared with a normally distributed random variable whose mean equals zero and standard deviation equals one. If the index value is less than or equal to the random variable value, the family is predicted to incur child-care expenses. If the index value is greater than the random variable value, the family is predicted not to incur child-care expenses (see Appendix A).

Based on this analysis, the characteristics which have an impact on whether a family incurs child-care expenses include:

- The number and age of ***children*** in the family
- The number of persons ***in the family,*** the number of ***families in the household,*** and the number of ***families in the household eligible for child care expenses***
- The demographic characteristics of the ***parent*** and the spouse
- The ***income*** and ***rental status*** of the family

We did not use variables that represent the receipt of public assistance or food stamps because the equations are also part of the process to simulate eligibility for other these programs. We discuss each of the characteristics included in the equation below.

1. The Number and Age of Children in the Family

As shown in Table II.2, the number and age of the children in a family are highly significant determinants of whether a family incurs child-care expenses. The probability of a family's incurring child-care expenses increases with the number of young children (younger than age 8), and decreases with the number of old children (older than age 11). The probability of a family's incurring child-care expenses increases with the average age of children younger than age 8.

The two most significant determinants of whether a family incurs child care expenses due to the number and age of children in the family are (1) the number of children in the family younger than age 5 (NCHLD4), and (2) the average age of the children in the family younger than age 5 (AAGECLD4); both variables are significant at the 1 percent level. The number of children in the family younger than age 5 *increases* the probability of a family's incurring child-care expenses, most likely because child-care services are especially needed for younger children who are not in school. The probability of a family's incurring child-care expenses increases with the presence and age of children younger than age 5, which is reasonable, since the mothers of infants are less likely to work than those with older toddlers. The number of relatively young children in the family, age 5 to 7, also increases the probability of a family's incurring child-care expenses (NCH5T7), as does the average age of these children (AAC5T7). It is interesting to note that the number of children age 8 to 11 was not a significant factor and was thus omitted from the equation; this insignificance suggests that finding care for children age 8 to 11 is less difficult than for younger children, or perhaps that these children may care for themselves.

The number of children in the family age 12 to 14 and 15 to 18 *decreases* the probability of a family's incurring child-care expenses (NCH12T14 and NCH15T18), most likely because, theoretically, older children are more capable of caring for themselves and their younger siblings. These variables are significant at the 1 percent level.

2. The Number of Persons in the Family, the Number of Families in the Household, and the Number of Families in the Household "Eligible" for Child-Care Expenses

As shown in Table II.2, an increase in the number of persons in a family (FNUMPER4) and the number of families in a household (HNF4) reduces the probability of a family's incurring child-care expenses, most likely because more caretakers are available to care for younger children without incurring an expense. (These variables are significant at the 1 percent level). However, as the number of families in a household eligible to answer the child care questions (NUMSFWCH) increases, so does the probability of a family's incurring child-care expenses (significant at the 11 percent level).

3. The Characteristics of the Parent and Spouse

The characteristics of the parent and spouse of a family are significant determinants of whether the family incurs child-care expenses. Specifically, we found that the age, race, and gender are significant factors. Also significant are the marital, educational, and work status of the parent, and the work status of the spouse. In general, the probability of a family's incurring child-care expenses increases if the family contains a single parent or two working parents, or if the parent works full-time. The probability of a family's incurring child-care expenses decreases as the parent's age increases.

More specifically, the probability of a family's incurring child-care expenses increases if the parent of the children is female or single; these variables (FEMALE and SINGLE1) are significant at the 1 and 5 percent levels, respectively. This finding supports the theory that single working parents have a greater need for child-care services, since they cannot share child-care responsibilities with another parent. In addition, the presence of a female parent increases the probability of a family's incurring child-care expenses. Similarly, if both the parent and spouse are working (EMPSTAT), the probability of the family's incurring child-care expenses increases, since neither parent is available to care for the children. The probability of a family's having positive child-care

expenses also increases with an increase in the usual number of hours a parent works per week (HRSWK), and if the parent works full-time as opposed to part-time (FULLPT). However, if a parent is self-employed (SELEMPL), the probability of a family's incurring child-care expenses decreases, most likely because self-employed parents may work at home and have a flexible work schedule.

The age of the parent is a highly significant determinant of whether a family incurs child-care expenses. The probability of a family's incurring child-care expenses decreases as the age of the parent (AGE4) increases. This relationship reflects the fact that older parents are less likely to have very young children, who are more likely to need child-care services. However, positive coefficients on the two "dummy" age variables--age 35 to 44 (AGE35T44) and older than age 44 (AGEGT44)--somewhat offset the effect of the age variable on the probability of a family's incurring child-care expenses. The combined effect of these variables indicates that, although the probability of a family's incurring child-care expenses declines with age, it declines in a nonlinear way.

To illustrate this relationship, we can examine the impact of the parent's age on the probability of a family's incurring child-care expenses, assuming that all else except age is held constant at the sample mean values and the dummy variable is excluded from the equations. For a family with a 20-year-old parent, the probability of the family's incurring child-care expenses is 33.1 percent. For a family with a 30-year-old parent, the probability of the family's incurring child-care expenses is 18.0 percent. For a family with a 40-year-old parent, the probability of the family's incurring child-care expenses based on the age variable only (AGE4) is 8.2 percent. However, if we add the effect of the dummy variable (AGE35T44) to the equation, the probability that a family with a 40-year-old parent incurs child-care expenses is 13.0 percent. Similarly, for a family with a 50-year-old parent, adding the impact of the dummy variable's (AGEGT44) lessens the decline in the probability of the family's incurring child-care expenses based only on the age variable.

Finally, the race and education of a parent are significant determinants of whether a family incurs child-care expenses. The probability of a family's incurring child-care expenses increases if the parent is nonwhite (NONWHITE); this variable is significant at the 16 percent level. The probability of a family's incurring child-care expenses increases as the parent's total number of years of education (HIGRADE) increases. However, two other education variables (parent attended some high school, and parent attended some college--SOMEHS and POSTHS) have negative coefficients, thus dampening the increase in the probability of a family's incurring child-care expenses as the total number of years of education completed by the parent increases.

4. **Family Income and Rental Status**

Several income-related characteristics and whether a family rents or owns its home have an impact on the probability of a family's incurring child-care expenses. The income variables that are determinants of the child-care expense status of a family include the log of the parent's hourly earnings (LNEARNPH) and the log of the combined earned and unearned income of the parent and spouse (LNTEARN and LNTUNERN). The probability that a family incurs child-care expenses increases as the log of the parent's earnings per hour and total earned and unearned income increase, most likely because the family has more income to spend on child-care expenses and because the opportunity cost of not working is higher. The probability that a family incurs child-care expenses decreases if the family rents a home (IRENT).

C. THECHILD-CAREEXPENDITUREEQUATION

The amount of the child-care expense incurred by a family is also a function of various demographic and economic characteristics of the family. The estimated expenditure equation approximates this function. Theoretical reasoning and prediction results were used to determine the final set of explanatory variables for this equation. The dependent variable is the amount of child-care expenses, and is estimated only for families whom the probability equations predicts to have

expenses. The expenditure equation also contains a random error term that is selected from a normal distribution whose mean equals to zero and whose standard deviation equals 0.551292 (see Appendix A).

The characteristics which have an impact on the amount of child-care expenses incurred by the family are presented in Table II.3 and include:

- The number and age of *children in* the family
- The age, race, and gender and the work, marital, and educational status of the parent, and the educational status of the spouse
- The *income* of the family members

Each of these characteristics is discussed below.

1. The Number of Children in the Family

As shown in Table II.3, the number of children in the family younger than age 5 (NCHLD4), age 5 to 11 (NCH5T11), and age 12 to 18 (NCH12T18) are significant determinants of the amount of child-care expenses incurred by a family. Similar to the theory underlying the probability equation, child-care expenses increase with the number of younger children, and child-care expenses decrease with the number of older children, who theoretically need less care because they are more likely to be in school. The variables representing the number of children younger than age 5 and age 12 to 18 are significant at the 1 percent level.

2. The Characteristics of the Parent and Spouse

The characteristics of the parent and spouse of the family are significant determinants of the child-care expense amount, just as they were in the determination of whether a family incurs child-care expenses. Specifically, the age, race, and gender and the work, marital, and educational status of the parent and the educational status of the spouse are significant factors affecting the amount of monthly child-care expenses.

The amount of a family's child-care expenses decreases if the parent is a single female (FEMSING), if the parent is nonwhite (NONWHITE), and if the parent is younger than age 25 (AGELT25). Although the first two characteristics increase the *probability* that a family incurs child-care expenses, these families spend less on child care than do other families. Similarly, the more hours a parent parent's educational level (HIGRADE) increases the child-care expense amount, the spouse's educational level reduces the amount.

3. **Family Income and Rental Status**

The per hourly wages of the parent and the spouse and the log of total unearned income have a significant impact on the amount of a family's child-care expenses. The amount of child-care expenses increases with each of these variables (EARNPHR1, SEARNPHR, AND LNTUNERN), which are highly significant (at the 1 and 5 percent levels). These findings imply that higher income-families spend more on child care. Finally, families that rent homes (IRENT) tend to spend more on child care than do homeowners.

III. EVALUATION OF THE CHILD-CARE EXPENSE EQUATIONS

The child-care expense equations estimated with SIPP data for the MATH model *predict* that a certain number of families have child-care expenses, and the amount of child-care expenses for those families. In this chapter, we evaluate how well the equations predicted these characteristics by comparing the predicted outcomes with reported outcomes. After describing our evaluation methodology, we discuss the results of our analysis for the probability and expenditure equations, and we compare the characteristics of families that report child-care expenses with those predicted to incur child-care expenses. In Chapter IV, we describe how well the equations perform when incorporated into the MATH model.

A. EVALUATION METHODOLOGY

We applied the child-care equations to the 1984 SIPP Panel, Wave 5, data with which they were estimated and compared predicted outcomes with reported outcomes among the following three groups of families:

- All families in the analysis file with which the equations were estimated, which includes families whose household income is less than or equal to 300 percent of the poverty level, referred to as "low-income families"
- Families whose household income is less than or equal to 130 percent of the poverty level, referred to as "potentially eligible families," since their incomes make them potentially eligible for the FSP
- Families in which at least one person reported food stamp receipt, referred to as "food stamp families"

For each of the groups of families, we evaluated how well the equations performed by comparing predicted outcomes with reported outcomes in SIPP. Specifically, to evaluate the probability equation, we compared the percentage of families predicted to incur child-care expenses with the percentage reporting child-care expenses. In addition, we examined the percentage of families whom

the equation "correctly" predicted to incur child-care expenses; that is, among the families predicted to incur child care expenses, we determined the percentage which actually reported child-care expenses. We also compared various economic and demographic characteristics of the families predicted to incur child-care expenses with those of families that reported child-care expenses. This comparison determines whether the families that incur child-care expenses according to the equation are similar to those that report child-care expenses. To evaluate the expenditure equation, we compared the average monthly child-care expense and the distribution of child-care expenses among families predicted to incur child-care expenses with those of families that reported child-care expenses.

B. THE CHILD-CARE PROBABILITY EQUATION

Using the probability equation, we estimated the number of families that incur child-care expenses.[1] Among low-income families, which is the universe for which the equations were estimated, 27 percent reported child-care expenses, and the equation predicted that 28 percent incur child-care expenses (Table III.1). The difference between reported and predicted outcomes is 1 percentage point, or 4 percent of those reporting an expense. Therefore, for low-income families, the probability equation closely predicts the percentage of families that incur child-care expenses.

The equations less accurately predicted the percentage of food stamp families and potentially eligible families to incur child-care expenses. The equations were less accurate for these subgroups because they do not contain as many families as does the low-income family group, and because they are a subset of the universe for which the equations were estimated. This finding suggests that these

[1] We applied the probability equation as described in Chapter II, including the random component. To evaluate the effectiveness of the random component, we also applied an equation which compared the index to a cutoff value instead of a normally distributed random number. Compared with using a cutoff value, the randomization method improves the accuracy of the equation's prediction. Specifically, 27.3 percent of low-income families reported child-care expenses. The equation with the random component predicted 27.7 percent of low-income families to incur child-care expenses (a difference of 2 percent), and the equation with the cutoff value predicted 23.3 percent of low-income families to incur child-care expenses (a difference of 15 percent).

TABLE 111.1

COMPARISON OF CHILD-CARE EXPENSE EQUATION RESULTS WITH REPORTED INFORMATION FOR LOW-INCOME, FOOD STAMP, AND POTENTIALLY ELIGIBLE FAMILIES

(1984 Panel of SIPP, Wave 5)

	Reported Data		Predicted Results		Predicted Correctly	
	Thousands	Percent	Thousands	Percent	Thousands	Percent
Low-Income Families						
Positive Child-Care Expenses	2,546	27	2,590	28	1,313	52
No Child-Care Expenses	6,791	73	6,747	72	5,514	81
Total	9,337	100	9,337	100	6,827	73
Average Child-Care Expenses	$151	NA	$150	NA	NA	NA
Food Stamp Families						
Positive Child-Care Expenses	212	25	257	30	111	52
No Child-Care Expenses	633	7s	589	70	489	77
Total	845	100	846	100	600	71
Average Child-Care Expenses	$104	NA	$128	NA	NA	NA
Potentially Eligible Families						
Positive Child-Care Expenses	512	22	575	25	224	44
No Child-Care Expenses	1,764	78	1,701	75	1,413	80
Total	2,276	100	2,276	100	1,637	72
Average Child-Care Expenses	$124	NA	$131	NA	NA	NA

SOURCE: 1984 Panel of SIPP, Wave 5.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level. Food stamp families are families in which at least one person receives food stamps. Potentially eligiile families are families whose household income is less than or equal to 130 percent of the poverty level.

NA = not applicable.

subgroups behave differently than the overall low-income population. Among food stamp families, 25 percent reported child-care expenses, and the equation predicted that 30 percent incur expenses, a difference of 5 percentage points, or 20 percent. Among potentially eligible families, 22 percent reported child-care expenses, and the equation predicted 25 percent, a difference of 3 percentage points, or 14 percent.

Although the equation may closely predict the total percentage of low-income families that incur child-care expenses, it may not predict the correct (or same) low-income families that reported child-care expenses. For 73 percent of the low-income families, the equation correctly predicted these low-income families to have or not to have child-care expenses (Table III.1). Of the low-income families predicted not to incur child-care expenses, 81 percent were predicted correctly, compared with 52 percent of those predicted to incur child-care expenses.

Although half of the low-income families predicted to incur child-care expenses did not report child-care expenses, this proportion is twice as high as the proportion which would be predicted correctly if these low-income families were assigned randomly to incur or not to incur child- care expenses in the absence of equations or other methods. Similarly, assigning low-income families randomly would yield an overall correctly predicted rate of 60 percent. By using these equations, we increased this percent to 73 percent. Therefore, compared with random assignment, these equations performed well at predicting low-income families to incur child-care expenses. The equation also performed well for food stamp and potentially eligible families; 71 percent and 72 percent respectively of these families were predicted correctly.

C. THE CHILD-CARE EXPENDITURE EQUATION

As shown in Table III.1, the equation accurately predicted average monthly child-care expenses for low-income families, but less accurately for food stamp and potentially eligible families. Among low-income families which reported child-care expenses, the average monthly child-care expense was $151. Among those predicted to incur child-care expenses, the child-care expense equation predicted

an average monthly child-care expense of $150, a difference of less than one percent. Among potentially eligible families, the difference between the reported and predicted average monthly child-care expense is $7, or 5 percent ($124, compared with $131). Among food stamp families, the equation is less accurate, overestimating the average monthly child-care expense by $24, or 23 percent ($104, compared with $128).

The distribution of each group of families that reported child-care expenses and were predicted to have child-care expenses by the amount of their child-care expenses are presented in Figure III.1 (low-income families), Figure III.2 (potentially eligible families), and Figure III.3 (food stamp families). The data upon which these figures are based are contained in Appendix C. For all low-income families, the equations overestimated the percentage of families that incur relatively low child-care expenses (from $51 to $150) and underestimated the percentage that incur mid-range child-care expenses (from $151 to $250). These equations also slightly overestimated the percentage of families that incur relatively high child-care expenses, and the average predicted monthly child-care expenses are higher than the reported expenses. This pattern of overestimating relatively low and relatively high child-care expenses, and underestimating child-care expenses in the middle ranges exists for potentially eligible and food stamp families as well. In addition, the shape of the pattern is different for these subgroups.

The discrepancy between the distribution of low-income families by reported and predicted child-care expenses may reflect nonlinear relationships between the explanatory and dependent variables in the child-care expense equations. The equations were based on a linear regression model and thus they do not capture nonlinear relationship exactly. These nonlinear relationships may contribute to the discrepancy between the distributions of families by reported and predicted child-care expenses for the subgroups as well; however, the discrepancies also reflect differences in the behavior of the subgroups compared with the overall population.

FIGURE 111.1

Distribution of Low-Income Families By
Monthly Child-Care Expenditures
1984 Panel of SIPP, Wave 5

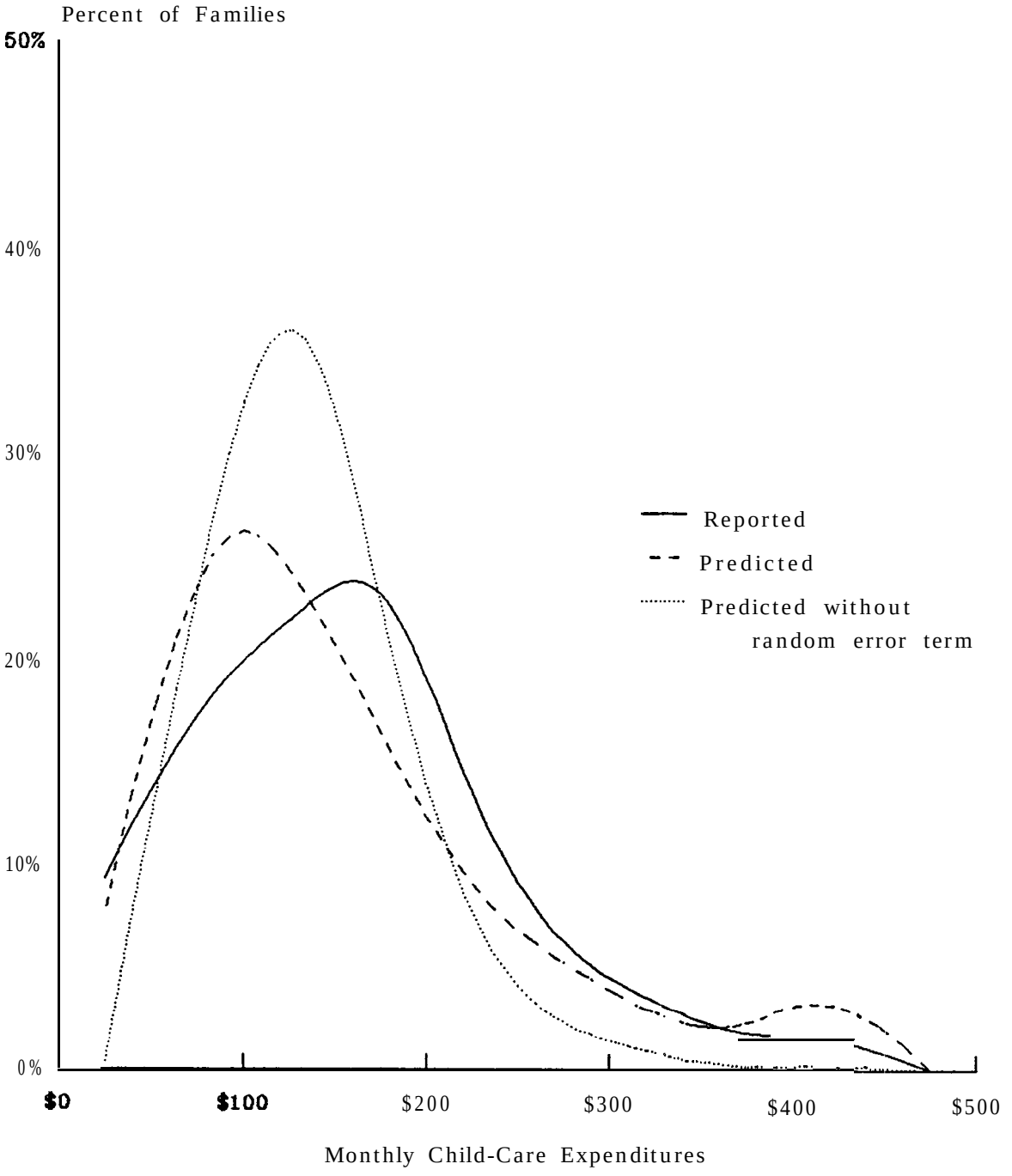


FIGURE III.2

Distribution of Potentially Eligible Families by Child-Care Expenditures

1984 Panel of SIPP, Wave **S**

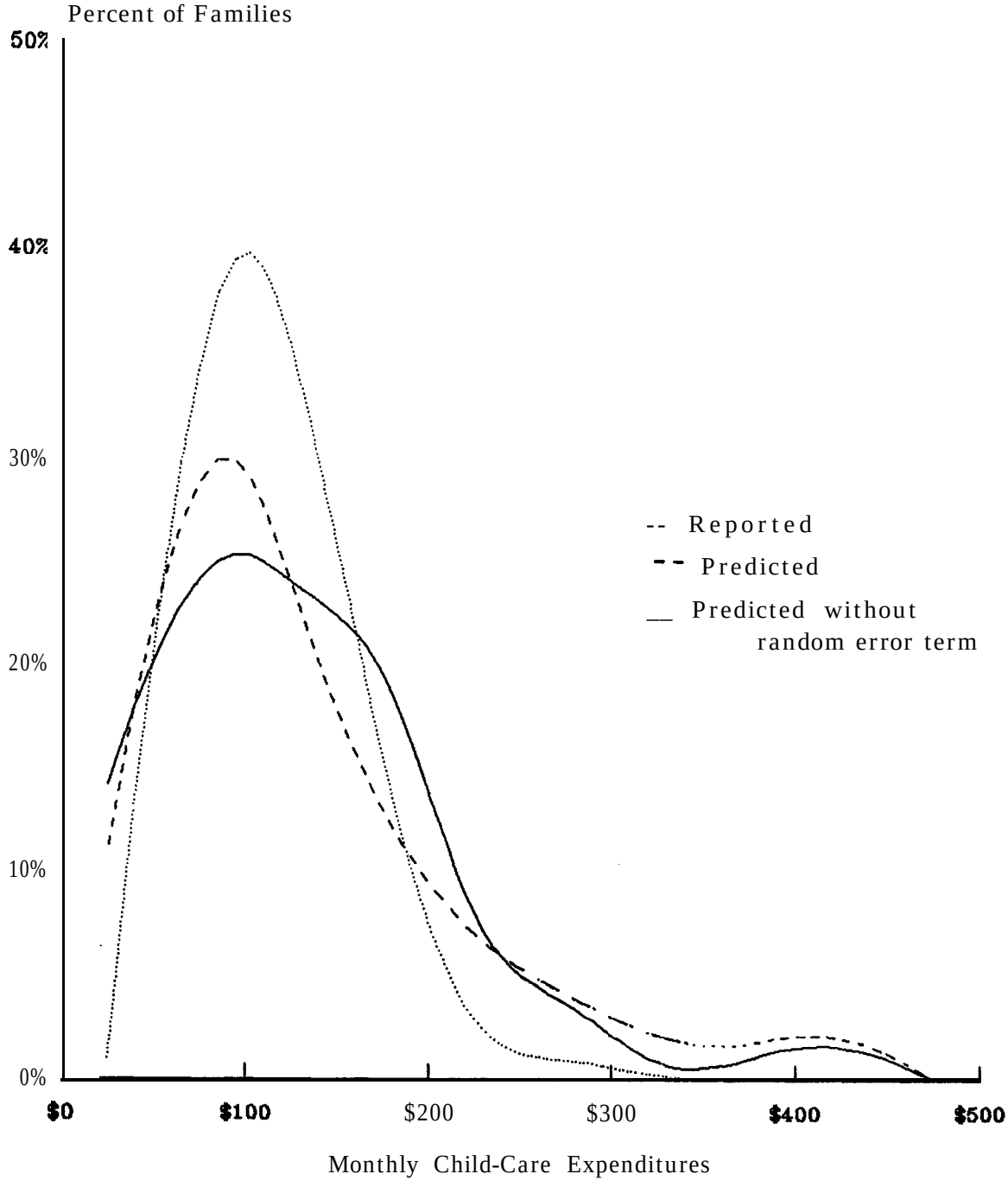
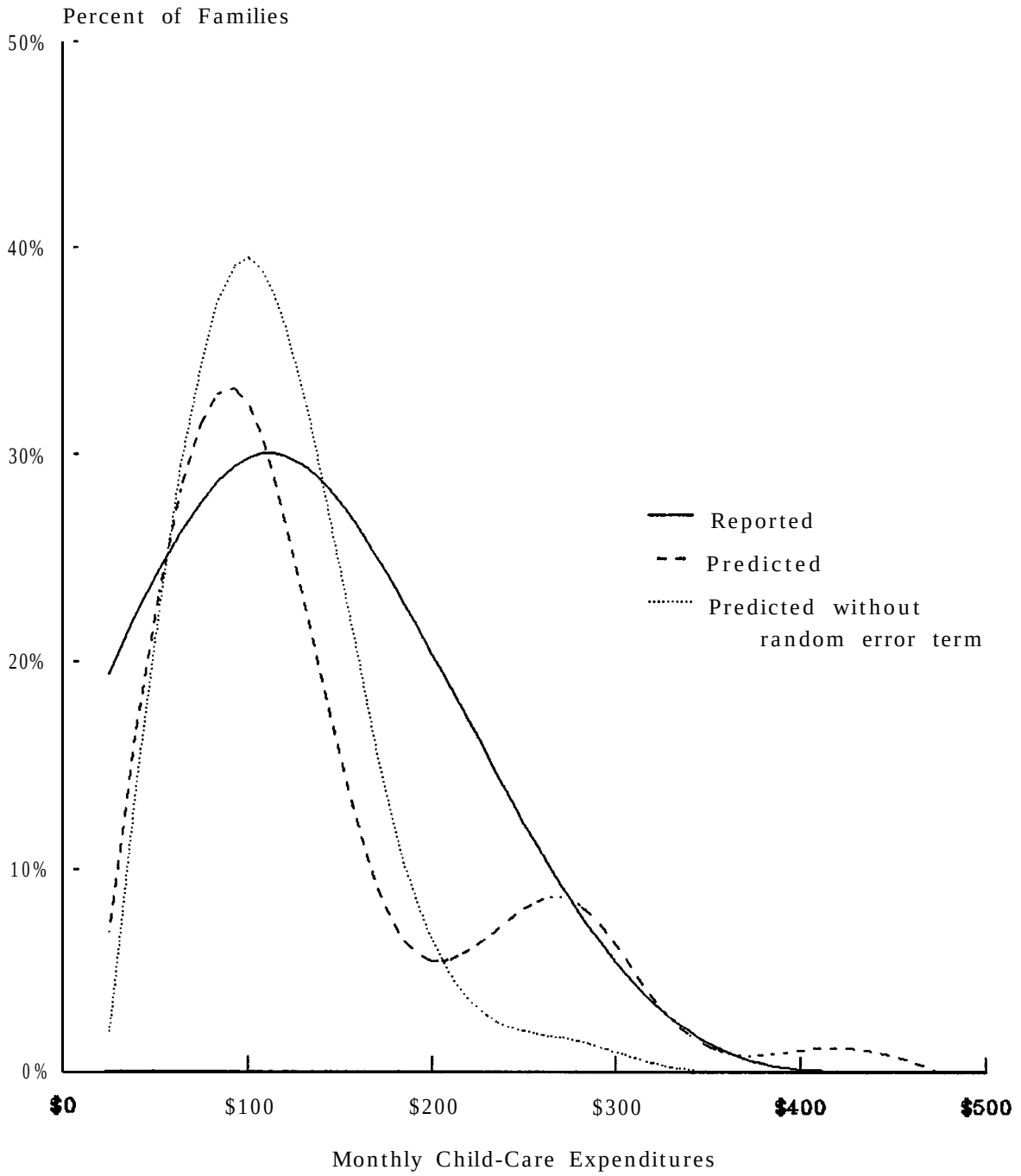


FIGURE III. 3

Distribution of Food Stamp Families By
Monthly Child-Care Expenditures
1984 Panel of SIPP, Wave 5



The expenditure equation also contains a random error term. To evaluate the impact of the random error term, we applied the expenditure equation without the error term to the SIPP data. As shown in Table III.2, the random component greatly improves the accuracy of the expenditure equation at estimating monthly child-care expenses for low-income families. Low-income families reported an average monthly child-care expense of $150. The equation with the random error term predicted $151. The equation without the random error term predicted $134, a difference of 11 percent. *Similarly, among* potentially eligible families, the random error term improved the accuracy of the prediction. Among food stamp families, however, the random error term reduced the accuracy. Food stamp families reported average monthly child-care expenses of $104, the equation with the random error term predicted $131, and the equation without the random error term predicted $109. As shown in the Figures the equation with the random error term performed better than the equation without the random error term.

D. DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS

As discussed earlier, the equation ***predicts*** that certain low-income families incur child-care expenses. Fifty-two percent of those families that were predicted to incur child-care expenses reported expenses. Therefore, in this section, we compare the demographic and economic characteristics of the low-income families that reported child-care expenses with those low-income families predicted to incur child-care expenses. We also make this comparison for food stamp and potentially eligiile families.

Selected demographic and economic characteristics for low-income families, potentially eligible. families, and food stamp families are shown in Table III.3. Among low-income families for which the equations were estimated, the characteristics of low-income families that reported child-care expenses and low-income families predicted to incur child-care expenses are very similar. For example, the average number of children in low-income families reporting child-care expenses is 1.8 persons, and

TABLE III.2

COMPARISON OF EXPENDITURE EQUATION RESULTS WITH AND WITHOUT A RANDOM ERROR TERM

(1984 Panel of SIPP, Wave 5)

	Average Monthly Child-Care Expenses		
	Reported	Predicted with Error Term	Predicted without Error Term
Low-Income Families	$151	$150	$134
Food Stamp Families	104	128	109
Potentially Eligible Families	124	131	109

SOURCE: 1984 Panel of SIPP, Wave 5.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level. Food stamp families are families in which at least one person receives food stamps. Potentially eligible families are families whose household income is less than or equal to 130 percent of the poverty level.

TABLE III.3

THE CHARACTERISTICS OF POTENTIALLY ELIGIBLE FAMILIES, FOOD STAMP FAMILIES, AND LOW-INCOME FAMILIES REPORTING CHILD-CARE EXPENSES AND PREDICTED TO HAVE CHILD-CARE EXPENSES

	Low-Income Families				Food Stamp Families				Potentially Eligible Families			
	Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)		Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)		Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Average Values:												
Average Monthly Child-Care Expenses	$151	n/a	$150	n/a	$104	n/a	$128	n/a	$124	n/a	$131	n/a
Average Number of Children In Family	1.8	n/a	1.9	n/a	1.8	n/a	2.2	n/a	1.9	n/a	2.1	n/a
Average Number of Persons in Family	3.7	n/a	3.7	n/a	3.2	n/a	3.7	n/a	3.7	n/a	3.8	n/a
Average Number of Persons in Household	3.8	n/a	3.8	n/a	3.3	n/a	3.9	n/a	3.8	n/a	3.9	n/a
Average Number of Families In Household	1.1	n/a	1.1	n/a	1.2	n/a	1.1	n/a	1.2	n/a	1.1	n/a
Average Number of Eligible Families in Household	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a
Average Age of Parent	28.8	n/a	28.7	n/a	26.4	n/a	26.4	n/a	27.3	n/a	27.7	nla
Average Hourly Earned Income of Parent	$5.20	n/a	$5.26	n/a	$3.80	n/a	f3.99	n/a	$3.64	n/a	$3.97	n/a
Average Hourly Earned Income of Spouse	$3.43	n/a	f3.59	n/a	so.45	n/a	SO.52	n/a	$0.58	n/a	SO. 82	n/a
Average Total Earned Income (Parent & Spouse)	$1,378	n/a	$1,383	n/a	$546	n/a	$519	n/a	$572	n/a	$608	n/a
Average Total Unearned Income (Parent & Spouse)	$97	n/a	$76	n/a	$26	n/a	$41	n/a	$41	n/a	$57	n/a
Average Hours Worked Per Week by Parent	35	n/a	35	n/a	31	n/a	32	n/a	32	n/a	33	n/e
Average Highest Grade Completed by Parent	13	n/a	13	n/a	12	n/a	11	n/a	12	n/a	12	n/a
Distribution of Families:												
Single Parent	1,022	40.2%	1,046	40.4%	177	83.6%	193	75.2%	324	63.4%	349	60.8%
Female Parent	2,485	97.6%	2,537	98.0%	199	94.0%	252	98.2%	494	96.5%	554	96.4%
Single Female Parent	973	38.2%	1,006	38.8%	lb4	77.6%	188	73.3%	312	61.0%	329	57.2%
Nonwhite Parent	639	25.1%	602	23.2%	66	31.0%	al	31.6%	171	33.4%	172	29.9%
Self-employed Parent	63	2.5%	58	2.2%	0	0.0%	0	0.0%	17	3.3%	9	1.6%
Parent Works Full-Tine	1,837	72.2%	1,826	70.5%	118	55.6%	143	55.8%	289	56.4%	346	60.1%
Parent & Spouse are Employed	2,386	93.7%	2,460	95.0%	198	93.5%	227	88.5%	451	88.1%	504	87.7%
Family Rents Shelter	1,194	46.9%	1,242	48.0%	159	75.2%	210	81.7%	345	67.5%	377	65.6%
Parent Attended Some College	866	34.0%	a35	32.3%	45	21.4%	39	15.2%	142	27.8%	124	21.5%
Parent Attended Some High School	1,609	63.2%	1,702	65.7%	167	78.6%	208	81.0%	357	69.8%	433	75.4%
Spouse Attended Some College	565	22.2%	601	23.2%	4	2.0%	4	1.7%	61	11.8%	46	8.0%
Spouse Attended Some High School	863	33.9%	a70	33.6%	30	14.4%	52	20.4%	109	21.2%	167	29.1%
All Families (Weighted in Thousands)	2,546	100.0%	2,590	100.0%	212	100.0%	257	100.0%	512	100.0%	575	100.0%

SOURCE: 1984 Panel of SIPP, Wave 5.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level.
Food stamp families are families in which at least one person receives food stamps.
Potentially eitgible families are families whose household income Is less than or equal to 130 percent of the poverty level.

the average number of children among low-income families predicted to incur child-care expenses is 1.9 persons.

However, the distribution and averages among food stamp families and potentially eligible families are less accurate than they are for low-income families. Specifically, among food stamp families, the equation predicted slightly larger families and households to incur child-care expenses. Food stamp families that reported child-care expenses have an average of 1.8 children, and the families that are predicted to incur child-care expenses have an average of 2.2 children. The equation also predicted fewer single-parent families (75 percent) compared with the percentage of single-parent families reporting child-care expenses (84 percent) to incur child-care expenses, a difference of 9 percentage points, or 12 percent. Finally, the equation does not accurately predict families to incur child-care expenses along the dimensions of the parent or spouse's educational status. For example, among food stamp families reporting child-care expenses, 21 percent contained a parent with some college, but the equation predicted only 15 percent.

The equations were developed for all low-income families and thus they perform well in selecting low-income families to incur child-care expenses which have characteristics similar to those who reported child-care expenses. Since the equations did not perform as well for the two subgroups, the behavior of these subgroups most likely differs from that of the overall population; thus, the estimated coefficients do not represent the relationship between the explanatory and dependent variables as accurately for the subgroups as they do for the overall population. To control for differences in the behavior of the potentially eligible subgroup, we can add a dummy variable and some interaction variables. However, if we include a dummy variable indicating food stamp families, imputed child-care expenses in the MATH database would depend on FSP participation modelled in the MATH model. Thus, we would not add a dummy variable for food stamp families.

IV. IMPLEMENTING THE CHILD-CARE EXPENSE EQUATIONS IN THE MATH MODEL

The child-care expense equations developed with SIPP data were implemented in the 1991 MATH model. Based on the FSP child-care deduction estimates which are calculated from the estimated expenses, we then used calibration factors to adjust the equations in order to meet the targeted dependent-care deduction amounts based on IQCS data. In this chapter, we describe the implementation and calibration processes and the calibration results.

A. THE IMPLEMENTATION PROCESS

Implementing the child-care expense equations into the MATH model entailed several steps. First, since income eligibility for food stamps is determined on a monthly basis and the CPS contains annual income data, the MATH model converted annual income data to monthly data. The child-care equations used the monthly income streams to estimate child-care expenses. Second, since the child-care expense equations were estimated at the family level (the most appropriate decision-making unit for child-care expenses), but the FSP dependent-care deduction applies to households, the model predicted family-level child-care expenses and then summed these to the household level in order to compute the child-care deduction. Third, since the equations were estimated with 1984 data, each of the equations was implemented in 1984 dollars (that is, all monetary independent variables were deflated to 1984 dollars prior to implementation); hence, predicted child-care expenses are in 1984 dollars. The model then used the overall Consumer Price Index for Wage Earners (CPI-W) to inflate the expenses predicted for the 1991 file to 1991. Finally, the model also simulated participation in public assistance programs other than the FSP, so the model incorporates the child-care equations in those simulations as well.

B. THE CALIBRATION PROCESS

After the child-care expense equations are applied to the MATH model, the resulting estimate of the dependent-care deduction may not closely match targeted amounts available from administrative (IQCS) data. Therefore, a calibration process must be implemented in order to adjust the predicted dependent-care deduction. The calibration process entails adjusting predicted child-care expenses; in doing so, the simulated dependent-care deduction for participating FSP households matches corresponding information derived from the most recent IQCS data. The process entails implementing additive and multiplicative adjustment factors to the equations and modifying these factors as necessary until the simulated deductions closely meet the targets. The adjustment process is iterative in the following way:

- Households from the CPS are selected to participate in the FSP based on the most recent participation probabilities.
- The model is executed with the most recent adjustment factors for the child-care expense equations, and simulated deductions are compared with observed deductions from the most recent IQCS data.
- Differences in deductions across the two data sources are used to compute new adjustment factors, and simulated deductions are compared with observed deductions.
- This process is repeated until differences are minimized.
- The model is executed after the participation algorithm is calibrated, and if the simulated deductions still do not meet the targets, the calibration process is repeated until differences are minimized.

The targeted amounts which govern this process are the average deduction for households with a dependent care deduction and the proportion of participating households with the dependent-care deduction. These values are estimated from the most recent IQCS file, in which the average amounts are projected forward (consistent with observed growth rates over the recent past) under the assumption that the proportion of households with the deduction remains constant over time. The 1991 control totals were estimated with IQCS data from summer 1988.

For the 1991 MATH model file, the process was modified slightly because the predicted 1991 child-care expense controls from the summer 1988 IQCS data did not reflect the legislated change in the dependent-care deduction cap from $160 per household to $160 per dependent in each household. Thus, we used an alternate food stamp model to modify the adjustment factors for child-care expenses. This alternate simulation agreed with the main model, except that it included the old rather than the new dependent-care deduction cap. (That is, it is customary to carry out the expense calibration using program regulations in effect during the simulation year. However, this is not advisable if there are significant changes in the eligibility criteria between the year of the most recent IQCS data and the simulation year.)

If the proportion of households with the deduction did not match the target, we adjusted the probability equation. If the proportion with expenses were on target but their average deduction was not on target, we then adjusted the expenditure equation. We adjusted average expenses only after the proportion was on target. If the proportion of households whose dependent-care deduction differed from the proportion in the IQCS data, we altered the equation adding a factor. Due to small sample sizes and the random component of the child-care estimation model, calibrating the proportion with child-care expenses often takes more than one attempt. To determine the adjustment factor for the expenses, we used the ratio of the target deduction to the average deduction from the MATH model.

C. THE CALIBRATION RESULTS

With the alternate 1991 model, the uncalibrated equations yielded an average deduction of $119.40 over 2.27 percent of participating households with the deductions. The projected targets were $115 and 2.8 percent. An additive factor of .18 and a multiplicative factor of .91 were applied, yielding an average deduction of $113.90 over 2.56 percent of the households. The additive factor was then increased to .255, yielding $113.89 and 2.77 percent. Then these adjustment factors were

used in the redetermination of food stamp eligibility under the 1991 regulation using the dependent-care deduction cap per dependent.

The initial adjustment of the participation probabilities adversely affected the child-care results. The proportion with the deduction declined to 2.32 percent, and the overall average deductions were too low. Thus, we increased the additive factor to .500, bringing the proportion with deductions up to 2.79 percent. (The average dependent-care deduction at this point was $122.79, reflecting the higher cap recently enacted, as well as the application of the adjusted participation probabilities.) After the participation algorithm was calibrated, the final average dependent-care deduction was $122.29, and 2.72 percent of participating households had the deduction, close to the targets of $115 and 2.8 percent.

V. EVALUATION OF THE CHILD-CARE EXPENSE EQUATIONS OVER TIME WITH THE 1987 SIPP PANEL, WAVE 6

The child-care expense equations were estimated with SIPP data from 1985 and were incorporated into the 1991 MATH model, which is based on the March 1988 CPS. To determine whether the child-care equations still reflect the relationship between family characteristics and their child-care expenses and can be again used in the next MATH model, we applied the child-care equations developed from the earlier SIPP data to the more recent SIPP data (the 1987 SIPP Panel, Wave 6) conducted in late 1988 and early 1989. Since the equations were developed with earlier data, we deflated the 1988 SIPP income data to 1984 dollars and then inflated them to 1988 dollars after estimating the child-care expenses.

In this chapter, we briefly describe the data analysis file and then discuss the results of applying the child-care probability and expenditure equations to the 1987 SIPP Panel, Wave 6. We applied two sets of equations, one that contained the calibration factors currently in place in the MATH model child-care expense equations, and one that did not. Our analysis entailed comparing reported and predicted outcomes among three groups of families: (1) low-income families, (2) potentially eligible families, and (3) food stamp families (as in Chapter III).

A. THE SIPP DATA ANALYSIS FILE

The SIPP child-care topical module survey questions changed between the 1984 SIPP Panel, Wave 5 and the SIPP 1987 Panel, Wave 6. The earlier SIPP survey asked child-care questions of working parents only. The more recent SIPP survey asked child-care questions of parents who were working, were looking for work, or were in school. In addition, the contents of the questions change between surveys. To be consistent with the data used to develop the child-care expense equations, we applied the child-care expense equations to the more recent SIPP data, restricting the universe to families with at least one working parent.

The difference between the new universe, consisting of working parents, parents looking for work or in school, and the old universe, consisting only of working parents, does not have a significant impact on our results. First, including newly eligible families in the universe increases the sample size only by 13 percent. Second, almost none of the newly eligible families had positive child-care expenses. The change in the questions asked may have an impact on this analysis because the accuracy of the data reported might have changed over time; however, we cannot separate this effect from changes in child-care usage over time.

On the 1987 SIPP Panel, Wave 6, 3,655 (unweighted) families had a working parent and children younger than 15. Screening out high-income families (those that belong to households whose income is greater than three times the poverty level) reduced the sample size to 2,007 families. These 2,007 families represent all families eligible to answer the child-care questions in the 1987 SIPP Panel, Wave 6, surveys. Since the child-care equations were developed based on the older universe, we screened out the newly eligible families, yielding a fmal sample size of 1,084 families (9.4 million families weighted). Among the 1,084 families, 352 (3.1 million weighted), or 33 percent, reported child-care expenses averaging $192 per month. Below, we evaluate how well the child-care expense equations predicted the probability of a family's incurring child- care expenses, and the amount of the child-care expenses.

B. THE CHILD-CARE PROBABILITY EQUATION

We applied the uncalibrated and calibrated child-care probability equations developed with the 1984 Panel of SIPP, Wave 5, data to the more recent SIPP data.[1] As shown in Table V.1, the uncalibrated equation predicted that 28 percent of low-income families incur child-care expenses, compared with 33 percent that reported child-care expenses, a difference of 5 percentage points, or 15 percent. Since the equations were estimated for low-income families and the calibration factors

[1]Since the calibration factors were developed to adjust the equations for the MATH model, we do not expect them to perform well but examine their impact on SIPP.

TABLE V.1

COMPARISON OF CHILD-CARE EQUATION EXPENSE RESULTS WITH REPORTED INFORMATION FOR LOW-INCOME, FOOD STAMP, AND POTENTIALLY ELIGIBLE FAMILIES

(1987 Panel of SIPP, Wave 6)

	Reported Data		Uncalibrated Predicted Results		Calibrated Predicted Results	
	Thousands	Percent	Thousands	Percent	Thousands	Percent
Low-Income Families						
Positive Child-Care Expenses	3,155	33	2,660	28	3,861	41
No Child-Care Expenses	6,276	67	6,770	72	5,569	59
Total	9,431	100	9,430	100	9,430	100
Average Child-Care Expenses	$192	NA	$175	NA	$157	NA
Food Stamp Families						
Positive Child-Care Expenses	228	33	252	37	323	47
No Child-Care Expenses	453	67	429	63	358	53
Total	681	100	681	100	681	100
Average Child-Care Expenses	$165	NA	$126	NA	$120	NA
Potentially Eligible Families						
Positive Child-Care Expenses	656	28	569	25	811	36
No Child-Care Expenses	1,664	72	1,751	75	1,508	65
Total	2,320	100	2,320	100	2,319	100
Average Child-Care expenses	$176	NA	$145	NA	$130	NA

SOURCE: 1987 Panel of SIPP, Wave 6.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level. Food stamp families are families in which at least one person receives food stamps. Potentially eligible families are families whose household income is less than or equal to 130 percent of the poverty level.

NOTE: Reported data are restricted to families with a working parent.
NA = not applicable.

were not developed for this universe, these results with the new SIPP data indicate that the uncalibrated equation is not accurate for the new data. This finding suggests that the relationship between the explanatory variables and the child-care expense status of the family has changed over time; indeed, there are substantial changes in characteristics of families reporting child-care expenses between the two surveys (Section D).

Similarly, among food stamp families, the calibrated equation predicted that 47 percent incur child-care expenses, compared with 33 percent reporting child care expenses, a difference of 14 percentage points, or 42 percent. Since the calibration factors were developed for food stamp families, these results are extremely inaccurate. In fact, for food stamp families, the uncalibrated equation performed better than the calibrated equation. However, the uncalibrated equation still overestimated the percentage of food stamp families that incur child-care expenses--by 5 percentage points, or 15 percent.

Neither equation performs well among potentially eligible families. The uncalibrated equation underestimated the percentage of families that incur child-care expenses--by 3 percentage points, or 11 percent--and the calibrated equation overestimated the percentage of families that incur child-care expenses--by 8 percentage points, or 29 percent.

In summary, for low-income families, the uncalibrated equation did not perform well and the calibrated equation performed even less accurately. For food stamp families, we would expect the cahbrated equation to perform well, but it did not; the uncalibrated equation performed better, but not satisfactorily. Among potentially eligiile families, the uncalibrated equation performed better than the calibrated equation, but still not satisfactorily. Differences in the characteristics of each group of families between the two surveys contribute to the performance of the child-care equation on the newer SIPP data.

In terms of the overall percentages of correctly predicted observations, the equation's performance compares favorably with its performance with the older SIPP data. As shown in Table

V.2, the uncalibrated equation correctly predicted between 69 and 74 percent of the families that have or do not have child-care expenses; the calibrated equation predicted between 68 and 72 percent correctly. As discussed in Chapter II, these proportions of the correctly predicted families are significantly better than simply assigning families randomly.

C. THE CHILD-CARE EXPENDITURE EQUATION

As shown in Table V.l, each of the equations (uncahbrated and calibrated) underestimated the average child-care expenses per month for each group of families, with the uncalibrated equation predicting more accurately in each case. Among low-income families, the uncalibrated equation predicted average monthly child-care expenses of $175, compared with reported expenses of $192--a difference of $17, or 9 percent. The discrepancy between reported and predicted average monthly child-care expenses with the uncalibrated equation for food stamp families and potentially eligiile families is $39 and $32, or 24 and 18 percent, respectively. (The calibrated equation prediction is inaccurate for each group of families.)

The distributions of each group of families by their reported and predicted child-care expenses are presented in Figure V.l (low-income families), Figure V.2 (potentially eligible families), and Figure V.3 (food stamp families). The data upon which these figures are based are contained in Appendix D. For all low-income families, the equation overestimated the percentage of families that incur relatively low child-care expenses (from $51 to $150) and underestimated the percentage that incur over $150. This pattern of overestimating relatively low child-care expenses and underestimating relatively high child-care expenses exists generally for potentially eligible and food stamp families, and is reflected in the lower average predicted child-care expenses compared with reported child-care expenses. As discussed below in section D, differences in the characteristics of families based on the two SIPP surveys explain, in part, why the equations predict lower child-care expenses compared with reported expenses.

TABLE V.2

CORRECTLY PREDICTED OBSERVATIONS WITH THE CHILD-CARE PROBABILITY EQUATION FOR LOW-INCOME, FOOD STAMP, AND POTENTIALLY ELIGIBLE FAMILIES

(1987 Panel of SIPP, Wave 6)

	Reported Data		Uncalibrated Predicted Correctly		Calibrated Predicted Correctly	
	Thousands	Percent	Thousands	Percent	Thousands	Percent
Low-Income Families						
Positive Child-Care Expenses	3,155	100	1,462	46	1,995	63
No Child-Care Expenses	6,276	100	5,078	81	4,410	70
Total	9,431	100	6,540	69	6,405	68
Food Stamp Families						
Positive Child-Care Expenses	228	100	143	63	162	71
No Child-Care Expenses	453	100	344	76	293	65
Total	681	100	487	72	455	67
Potentially Eligible Families						
Positive Child-Care Expenses	656	100	311	47	405	62
No Child-Care Expenses	1,664	100	1,406	84	1,258	76
Total	2,320	100	1,717	74	1,663	72

SOURCE: 1987 Panel of SIPP, Wave 6.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level. Food stamp families are families in which at least one person receives food stamps. Potentially eligible families are families whose household income is less than or equal to 130 percent of the poverty level.

NOTE: Reported data are restricted to families with a working parent.

FIGURE V.1

Distribution of Low-Income Families By Monthly Child-Care Expenditures

1987 Panel of SIPP, Wave 6

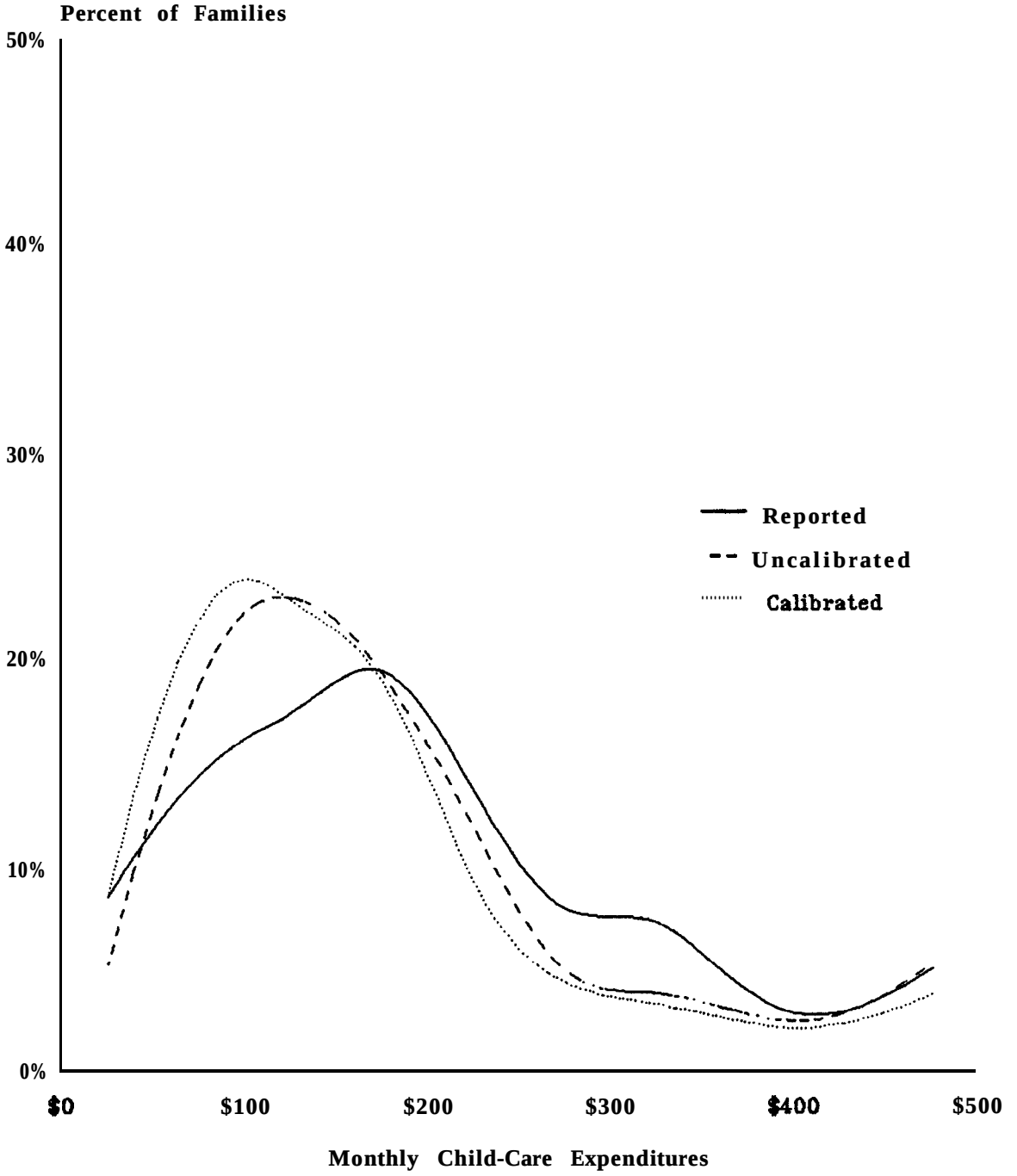


FIGURE V.2

Distribution of Potentially Eligible Families by Child-Care Expenditures

1987 Panel of SIPP, Wave 6

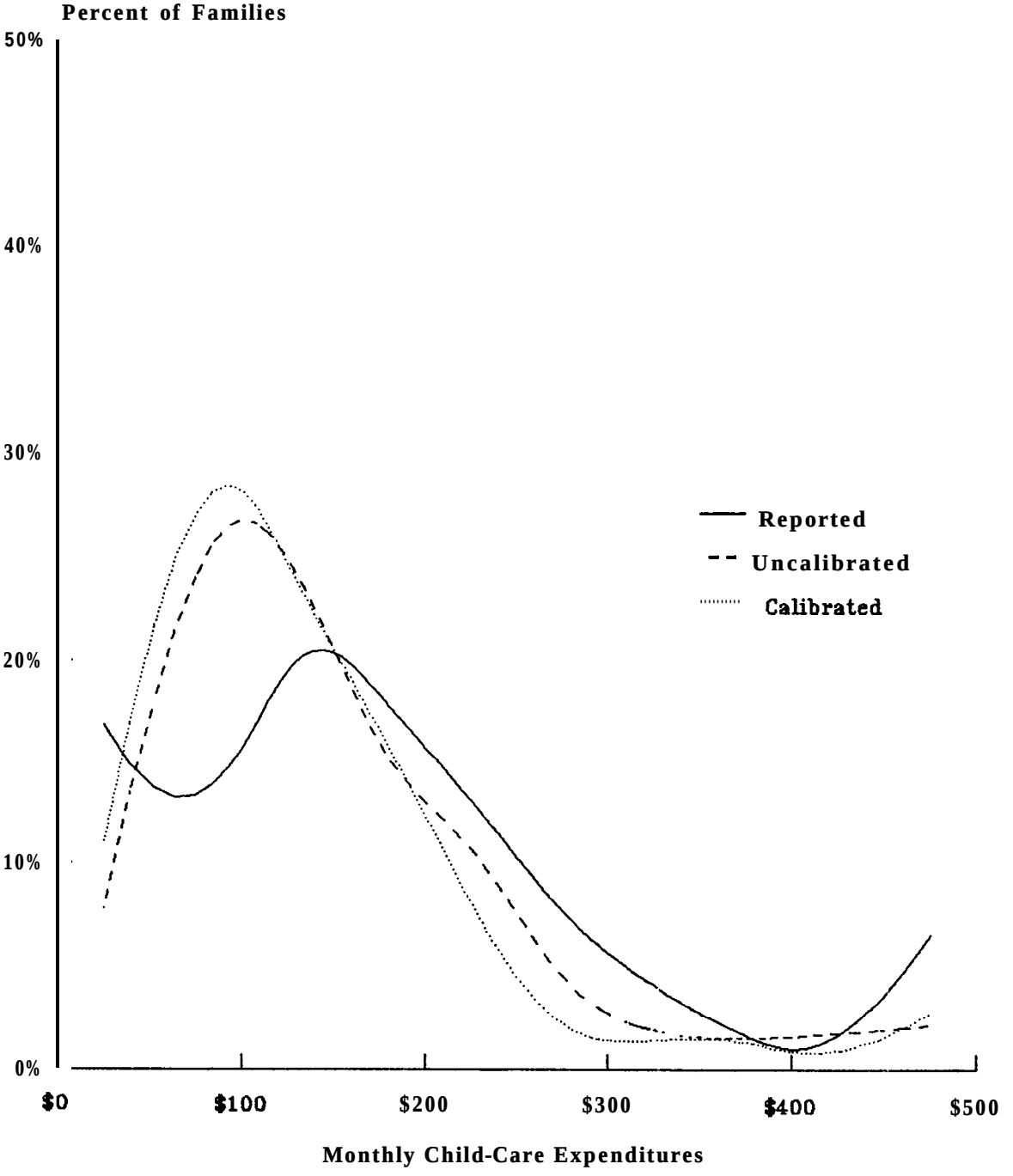
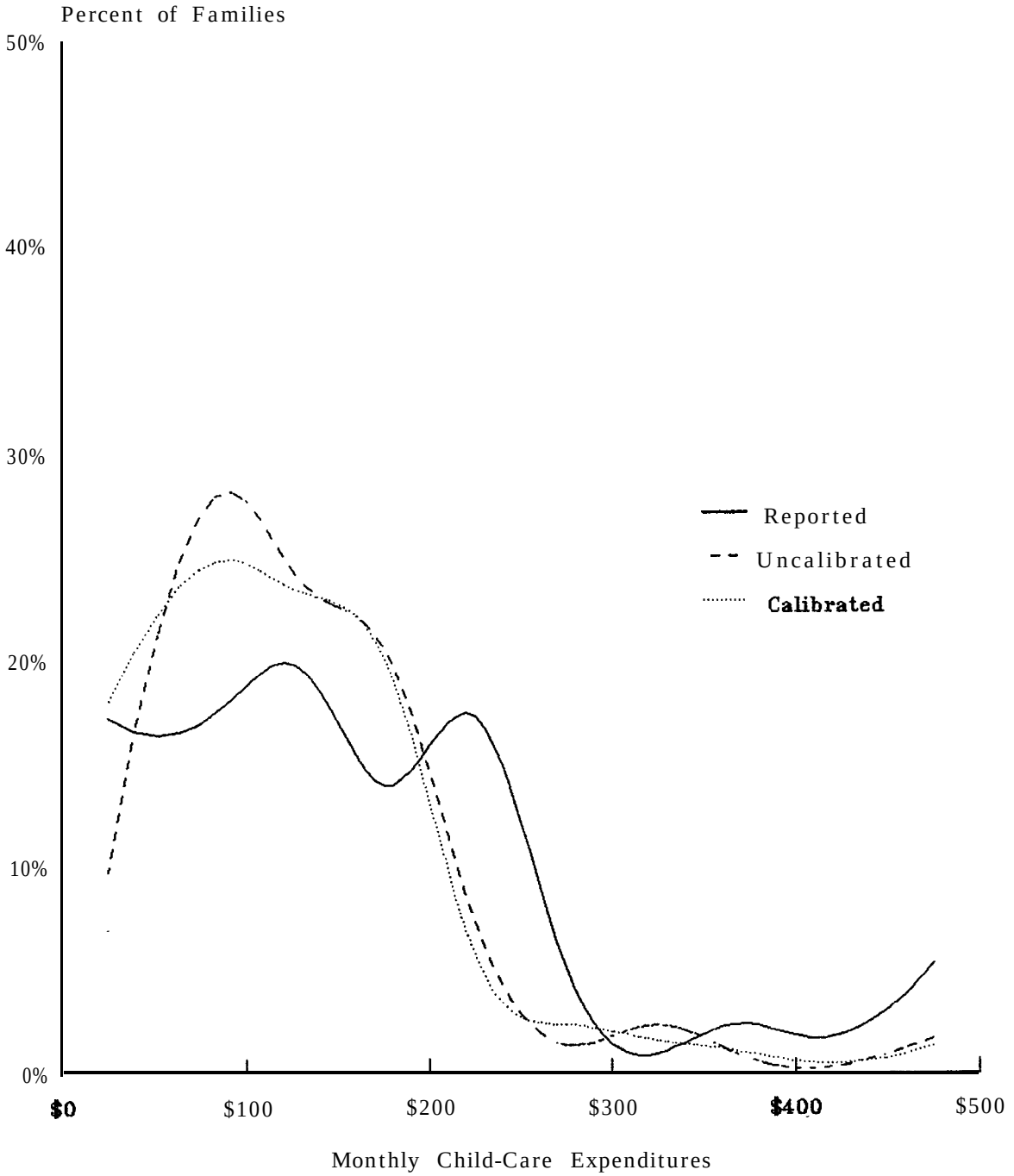


FIGURE V.3

Distribution of Food Stamp Families By
Monthly Child-Care Expenditures
1987 Panel of SIPP, Wave 6



D. DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS

As discussed earlier, the equation *predicts* that certain families to incur child-care expenses. Not all of these families reported child-care expenses. Therefore, in this section we compare the demographic and economic characteristics of the families that reported child-care expenses with those families whom the equation predicted incur child-care expenses. We also compare the characteristics of the families reporting child-care expenses based on the recent SIPP data to those reporting child-care expenses based on the 1984 SIPP Panel, Wave 5.

Selected demographic and economic characteristics for low-income families, potentially eligible families, and food stamp families are shown in Table V.3. For low-income families, the characteristics of families that reported child-care expenses and those of families predicted to incur child care expenses are very similar. However, both the uncalibrated and calibrated equations select families with slightly less total earned income and slightly more total unearned income to incur child-care expenses. In addition, both equations select more families that contain a nonwhite parent, and single female parents to incur child-care expenses.

These differences are more dramatic with potentially eligible and food stamp families. Among food stamp families, both equations select families with less total unearned income (reflecting the fact that unearned income is not a highly significant factor in the equation) to incur child-care expenses. More important, the equations select less families that contained single or single female parents, by more than 10 percent, and overestimated the percentage of families that contained a nonwhite parent, by over 10 percent to incur child-care expenses. These findings are generally the same for potentially eligible families. In addition, the equation did not accurately predict families to incur child-care expenses based on the parent's or spouse's educational status.

These differences in the characteristics of families reporting child-care expenses and those predicted to incur child-care expenses based on the 1987 SIPP Panel, Wave 6 are much larger than the differences between these two groups based on the 1984 SIPP Panel, Wave 5. This finding, as

TABLE V.3

THE CHARACTERISTICS OF POTENTIALLY ELIGIBLE FAMILIES, FOOD STAMP FAMILIES, AND LOW-INCOME FAMILIES REPORTING CHILD-CARE EXPENSES AND PREDICTED TO HAVE CHILD-CARE EXPENSES

	Low-Income Families						Food Stamp Families						Potentially Eligible Families					
	Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)		Families with Predicted Child-Care Expenses (Calibrated)		Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)		Families with Predicted Child-Care Expenses (Calibrated)		Families with Reported Child-Care Expenses		Families with Predicted Child-Care Expenses (Uncalibrated)		Families with Predicted Child-Care Expenses (Calibrated)	
Characteristic	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Average Values:																		
Average Nonthly Child-Care Expenses	$192	n/a	$175	n/a	$157	n/a	$165	n/a	$126	n/a	$120	n/a	$176	nla	$145	nla	$130	n/a
Average Number of Children In Family	2.0	n/a	1.9	nla	1.9	n/a	2.6	n/a	2.6	n/a	2.4	n/a	2.3	n/a	2.3	nla	2.1	n/a
Average Number of Persons in Family	3.7	n/a	3.7	n/a	3.6	n/a	4.1	n/a	4.1	n/a	3.9	n/a	3.8	nla	3.7	nla	3.6	n/a
Average Number of Persons in Household	3.9	n/a	4.0	nla	3.9	n/a	4.2	n/a	4.1	n/a	4.0	n/a	3.9	nla	3.9	n/a	3.8	n/a
Average Number of Families in Household	1.1	n/a	1.1	nla	1.1	n/a	1.1	n/a	1.0	n/a	1.1	n/a	1.1	nla	1.1	n/a	1.1	n/a
Average Number of Eligible Families in Household	1.0	n/a	1.1	n/a	1.1	n/a	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a	1.0	n/a
Average Age of Parent	29.1	n/a	29.1	n/a	29.9	n/a	29.1	n/a	29.2	n/a	29.8	n/a	29.1	n/a	28.8	n/a	29.4	n/a
Average Hourly Earned Income of Parent	$4.90	n/a	S4.75	nla	f4.77	n/a	$3.84	n/a	$3.38	n/a	$3.40	n/a	f3.49	n/a	$3.25	n/a	$3.27	n/a
Average Hourly Earned Income of Spouse	$3.93	n/a	S3.78	n/a	$3.70	n/a	$0.65	n/a	$0.72	n/a	$0.76	n/a	$0.65	n/a	$0.93	n/a	$0.81	n/a
Average Total Earned Income (Parent & Spouse)	$1,354	nla	$1,283	n/a	$1,299	n/a	$527	n/a	$552	n/a	$529	n/a	$554	nla	$550	nla	$532	n/a
Average Total Unearned Income (Parent & Spouse)	$68	n/a	$79	n/a	$80	n/a	$100	n/a	$55	n/a	$55	n/a	$87	nla	$56	n/a	$56	n/a
Average Hours Worked Per Week by Parent	35	n/a	35	n/a	35	n/a	32	n/a	34	n/a	32	n/a	35	nla	34	nla	34	n/a
Average Highest Grade Completed by Parent	12	n/a	12	n/a	13	n/a	11	n/a	12	n/a	12	n/a	12	nla	12	n/a	12	n/a
Distribution of Families:																		
Single Parent	1,260	39.9%	1,082	40.7%	1,611	41.7%	188	82.4%	174	68.9%	215	66.4%	471	71.8%	379	66.7%	545	67.2%
Female Parent	3,010	95.4%	2,606	98.0%	3,735	96.7%	228	100.0%	253	100.0%	323	100.0%	629	95.9%	564	99.2%	797	98.3%
Single Female Parent	1,115	35.4%	1,028	39.6%	1,485	38.5%	188	82.4%	174	68.9%	215	66.4%	444	67.7%	375	65.9%	531	65.4%
Nonwhite Parent	772	24.5%	798	30.0%	1,093	28.3%	95	41.8%	152	60.2%	179	55.4%	196	29.8%	264	46.4%	354	43.7%
Self-employed Parent	163	5.2%	138	5.2%	174	4.5%	12	5.3%	12	4.8%	12	3.8%	42	6.4%	38	6.6%	38	4.7%
Parent Works Full-Time	2,131	67.5%	1,876	70.5%	2,605	67.5%	134	58.9%	153	60.4%	175	54.1%	434	66.2%	339	59.6%	460	56.8%
Parent & Spouse are Employed	2,866	91.5%	2,479	93.2%	3,535	91.6%	216	94.7%	213	84.4%	258	79.9%	570	87.0%	507	89.2%	694	85.6%
Family Rents Shelter	1,623	51.5%	1,208	45.4%	1,764	45.7%	170	74.6%	165	65.2%	200	61.9%	443	67.5%	328	57.7%	501	61.8%
Parent Attended Some College	1,261	40.6%	908	34.1%	1,410	36.5%	75	32.7%	66	26.1%	76	23.6%	283	43.2%	191	33.6%	247	30.5%
Parent Attended Some High School	1,713	54.3%	1,633	61.4%	2,299	59.5%	111	48.6%	167	66.0%	227	70.2%	308	46.9%	351	61.7%	526	64.9%
Spouse Attended Some College	817	25.9%	573	21.5%	782	20.2%	20	8.6%	20	7.7%	24	7.5%	82	12.5%	61	10.7%	71	8.8%
Spouse Attended Some High School	992	31.5%	920	34.6%	1,356	35.1%	11	5.0%	31	12.1%	44	13.7%	84	12.8%	98	17.3%	149	18.3%
All Families (Weighted in Thousands)	3,154	100.0%	2,660	100.0%	3,861	100.0%	228	100.0%	253	100.0%	323	100.0%	656	100.0%	569	100.0%	811	100.0%

SOURCE: 1987 Panel of SIPP, Wave 6.

NOTE: Low-income families are families whose household income is less than or equal to 300 percent of the poverty level.
Food stamp families are families in which at least one person receives food stamps.
Potentially eligible families are families whose household income is less than or equal to 130 percent of the poverty level.

well as the finding that the equations predict lower child-care expenses compared with reported expenses, reflect in part differences between the families reporting child-care expenses in the two surveys. Compared with low-income families reporting child-care expenses in the 1984 SIPP Panel, Wave 5, low-income families reporting child-care expenses in the 1987 SIPP Panel, Wave 6 reported higher monthly child-care expenses, but lower total earned income and total unearned income, and lower hourly earned income. In addition, a lower proportion of families in the more recent SIPP data had a parent who worked full-time, or had both the parent and spouse employed. Given these characteristics--that families in the more recent SIPP survey were less likely to be employed full-time and had less earned and unearned income--it follows that our equation would predict that their child-care expenses would be lower than families in the earlier survey.' However, the reported child-care expenses were higher for families in the more recent SIPP data. Therefore, we conclude that the behavior of low-income families has changed between the two surveys, making the behavior captured in the child-care expense equations less accurate.

[2]Each of these variables has a positive relationship with the amount of child-care expenses.

VI. IMPLICATIONS OF THE EVALUATION OF THE CHILD-CARE EXPENSE EQUATIONS FOR THE MATH MODEL

The child-care expense equations evaluated in this report were incorporated into the 1991 MATH model to estimate child-care expenses and the dependent-care deduction for food stamp households, ultimately to support analyses of reforms to the FSP. Future MATH models based on the March CPS will also need child-care expense equations to estimate child-care expenses and the dependent-care deduction. In this chapter, we summarize the results of the evaluation of the child-care expense equations and discuss the implications of the results for the equations for future MATH models.

The evaluation of the child-care expense equations in this report consists of applying the equations to SIPP data and comparing the reported outcomes to the outcomes predicted by the equations. Specifically, the evaluation includes a comparison of predicted outcomes to reported outcomes based on the SIPP data from which the equations were developed (1984 SIPP Panel, Wave 5) and on more recent SIPP data (1987 SIPP Panel, Wave 6). The first comparison ensures that the equations used in the 1991 MATH model accurately predict child-care expenses for low-income families and the second comparison investigates whether the equations accurately predict child-care expenses for low-income families or more recent SIPP data. We also evaluated how well the calibrated equations performed for food stamp families using the more recent SIPP data, since the calibration factors were developed to improve the accuracy with which the equations estimate child-care expenses and the child-care deduction for food stamp households.

Based on the SIPP data upon which the equations were developed, we found that the equations accurately predicted child-care expenses for the universe they were estimated for, low-income families. After applying the uncalibrated equations to the more recent SIPP data, we found that the uncalibrated equations did not perform well for low-income families. This finding suggests that the relationship between the explanatory variables in the probability and expenditure equations and the

dependent variables has changed over time, or that some new factors not included in the equations are influencing the relationship. Differences in the characteristics of low-income families between the two surveys supports the notion that the behavior of low-income families changed over time.

We also found that potentially eligible and food stamp families behave differently from the overall population. Based on the evaluation, we found that the calibrated child-care equations did not perform well for food stamp families, even though the calibration factors were developed to adjust the child-care information so that the resulting estimated child-care deduction reflect targets from IQCS data. (The uncalibrated equations also did not perform well). This finding suggests that there may be an underlying discrepancy between the SIPP data and the administrative IQCS data from which we obtain the targets.

A comparison between the data sets confirms the notion that an underlying discrepancy between the two data sets contributes to the inadequacy of the equations. The average monthly child-care expenses and the percentage of food stamp families or food stamp households reporting child-care expenses based on IQCS and SIPP data for 1985 and 1988/1989 are presented in Table VI.1.[1] The average monthly child-care expenses are higher in SIPP than in the IQCS data and this difference increased dramatically from 1985 to 1988/1989 from 9 to 54 percent. Similarly, the percentage of food stamp families or households reporting child care expenses is significantly higher based on SIPP data than IQCS data. Based on the SIPP data, at least one-quarter of the food stamp families reported child-care expenses, compared to less than 3 percent of the food stamp households based on the IQCS data.

Since the equations did not perform well on the recent SIPP data and since there is a discrepancy between the IQCS and the SIPP data, we recommend the following activities.

[1]The SIPP data refers to families and the IQCS data refers to households.

TABLE VI.1

COMPARISON OF AVERAGE MONTHLY CHILD-CARE EXPENSES BASED ON IQCS AND SIPP DATA

	1984 Data	1988/1989 Data
Food Stamp Families (SIPP)		
Average Monthly Child-Care Expenses	$104	$165
Percentage with Child-Care Expenses	25%	33%
Food Stamp Households (IQCS)		
Average Monthly Child-Care Expenses	$95	$107
Percentage with Child-Care Expenses	1.8%	25%

SOURCE: 1984 Panel of SIPP, Wave 5, 1987 Panel of SIPP, Wave 6, Summer 1985 IQCS sample, and winter 1988, IQCS sample.

- ***We recommend investigating the discrepancy between the SIPP and IQCS data sets to determine if the relationship between the explanatory and dependent variables differ for the two data sets.*** One approach to investigating this discrepancy would be to apply the equations to the IQCS data and compare the resulting coefficients with those obtained from the SIPP data. However, the IQCS data do not contain sufficient information for the equations to operate. Thus, we recommend a less sophisticated alternative of comparing the relationship of various characteristics to child-care expenses in the IQCS to the relationship of the characteristics in SIPP. For example, we could examine the probability of a family reporting child-care expenses if a child under age 5 is present in each of the data sets.

- ***We recommend using the existing equations and reestimating the coefficients so that the equations represent more precisely the relationship between the explanatory and dependent variables. This*** reestimation would take into account the changes in behavior of low-income families over time, as well as any impact the change in SIPP's child-care questions may have had on the performance of the equations. Although we cannot be any more confident that updating to the 1987 SIPP Panel, Wave 6 will predict child-care expenses in a future year better than the existing equations, we would not recommend abandoning these estimation procedures based on the outcome of one application only, especially since the quality of the data may have changed between the surveys.

- ***If reestimating the existing equations on more recent SIPP data does not improve the results, we would recommend redesigning the modelling procedures to capture (1) differences in behavior among the potentially eligible group through the introduction of a dummy variable, and (2) nonlinear relationships between the explanatory and dependent variables.***

REFERENCES

Doyle, Pat, Judy Richter, Richard Shin, and Carole Trippe. "Creation of the 1991 MATH Data Base from the March 1988 Current Population Survey. Washington, DC: **Mathematica** Policy Research, Inc., 1991.

Doyle, Pat and Carole Trippe. "Improving the Income Allocation Procedures in MATH". Washington, DC: **Mathematica** Policy Research, Inc., 1991.

Heiser, Nancy. "Characteristics of Food Stamp Households: Summer 1989." Alexandria, VA: Food and Nutrition Service, U.S. Department of Agriculture, 1991.

Martini, Alberto. "Evaluating the Food Stamp Participation Algorithms used in Microsimulation Models." Washington, DC: **Mathematica** Policy Research, Inc., 1991.

Wonnacott, Thomas H. and Wonnacott, Ronald J. ***Introductory Statistics for Business and Economics: Third Edition. New*** York: John Wiley & Sons, 1984.

APPENDIX A

THE PROBABILITY AND EXPENDITURE EQUATIONS

In this appendix, we describe the child care probability and expenditure equations in more detail.

A. PROBABILITY EQUATION

The probability of a CPS family's incurring child-care expenses is a function of the various demographic and economic characteristics of the family. Based on this function, the MATH model calculates a value **(PCHILD)** 1 if the family incurs child-care expenses, and 0 otherwise. The value of PCHILD is a function of an index **value (CINDEX).** The value of CINDEX is compared with a randomly distriiuted variable to determine the value of PCHILD, as presented below:

PCHILD = 1 if $-\epsilon$ <= CINDEX
PCHILD = 0 if $-\epsilon$ > CINDEX

where:

$$\text{CINDEX} = \sum_{j=1}^{n} \beta_j X_j$$

where:

β_j	=	Coefficients for family characteristics variables.
4	=	The value of family characteristic variables.
ϵ	=	Normally distributed random variable with a mean of 0, and a standard deviation of 1.

B. EXPENDITURE EQUATION

The amount of the child-care expense incurred by a family is also a function of the various demographic and economic characteristics of the family. Based on this function, the MATH model predicts the amount of child-care expenses **(MNCHILD)** for families predicted to incur child-care expenses **(PCHILD=1),** as presented below:

MNCHILD = **[MIN (4.333*(EXP(LNMNCHLD))), 433] if PCHILD = 1.**

MNCHILD = 0, if **PCHILD** = 0.

where:

$$\text{LNMNCHLD} = \beta_1 + \sum_{j=1}^{n} \beta_j Z_j + \Phi.$$

β_1	=	Intercept term.
β_j	=	Coefficients for family characteristic variables.
Z_j	=	Value of family characteristics.
Φ	=	Random disturbance term that must be selected from a normal distriiution with a mean of 0 and a standard deviation of .5512920.

APPENDIX B

INTERPRETATION OF CHILD-CARE EXPENSE EQUATION RESULTS

The results of MPR's child-care equations are presented in Tables II.2 and lI.3. The variables in these tables either have a statistically significant impact on whether a family incurs child-care expenses or on the amount of child-care expenses, or is theoretically an important determinant.

The first column lists the variable name. (See Table II.1 for a more detailed description of the variables.). The column labeled "coefficient" lists the coefficients for each variable. If the coefficient for a given variable in the table is positive, the probability of a family's incurring child- care expenses or the amount of child-care expenses increases with the variable. Conversely, if the coefficient for a given variable in the table is negative, the probability of a family's incurring child- care expenses or the amount of child-care expenses decreases with the variable. For example, as shown in Table lI.2, the coefficient for the variable which represents the age of the parent (AGE4) is negative (--0477598); thus, as the age of the parent increases, the probability of a family's incurring child-care expenses decreases. In Table II.3, the coefficient that indicates the gender and marital status of a parent is a "dummy variable," which means that the value of the variable is one or zero. In this case, if the variable FEMSING equals one, the parent is a single female; if the variable equals zero, the parent is not a single female. Since the coefficient for the variable FEMSING is negative (-.310098), the presence of a single female parent reduces the amount of family child-care expenses relative to families whose parents are not single females.

The column labelled "signif. level" lists the significance level for each of the variables. For example, in Table II.2, the variable AGE4, which represents the age of the parent, is statistically significant at the 1 percent level. The variable that represents the usual number of hours that the parent works per week (HRSWK) is significant at the 16 percent level, which means that it is a less significant determinant of whether a family incurs child care expenses than is the age of the parent. For a more detailed discussion of significance testing, see Wonnacott and Wonnacott, 1984.

APPENDIX c

PERCENT DISTRIBUTION OF FAMILIES BY TYPE AND BY MONTHLY CHILD-CARE EXPENSES 1984 SIPP PANEL, WAVE 5

APPENDIX C

PERCENT DISTRIBUTION OF FAMILIES BY TYPE AND BY MONTHLY CHILD-CARE WPENSES
1984 SIPP PANEL, WAVE 5

Type of Family and Monthly Child-Care Expenses	Reported Child-Care Expenses	Predicted Child-Care Expenses	Change in Percentage Points
Low-Income Families			
l-50	**9.3**	7.6	-1.7
51-100	18.6	28.8	10.2
101-150	21.6	25.6	4.0
151-200	26.3	16.2	-10.1
201-250	12.3	a.2	-4.1
251-300	5.4	5.1	-0.3
301-350	3.3	2.5	-0.8
351-400	1.2	1.4	0.2
401-450	1.9	4.7	2.8
451+	0.0	0.0	0.0
Total	100.0	100.0	0.0
Potentially Eligible Families			
1-50	14.2	**11.0**	-3.2
51-100	28.1	37.6	9.5
101-150	23.1	22.6	-0.5
151-200	21.8	12.0	-9.8
201-250	5.3	6.4	1.1
251-300	4.3	4.2	-0.1
301-350	0.0	1.9	1.9
351-400	0.9	1.3	0.4
401-450	2.4	3.0	0.6
451+	0.0	0.0	0.0
Total	100.0	100.0	0.0
Food Stamp Families			
1-50	**19.4**	**6.9**	-12.5
51-100	**29.3**	42.6	13.3
101-150	24.5	26.0	1.5
151-200	16.2	4.1	-12.1
201-250	**8.1**	5.7	-2.4
251-300	2.s	10.9	a.4
301-350	0.0	1.9	1.9
351-400	0.0	0.0	0.0
401-450	0.0	1.9	1.9
451+	0.0	0.0	0.0
Total	100.0	100.0	0.0

SOURCE: 1984 Panel of SIPP, Wave 5.

APPENDIX D

PERCENT DISTRIBUTION OF FAMILIES BY TYPE
AND BY MONTHLY CHILD-CARE EXPENSES
1987 SIPP PANEL, WAVE 6

APPENDIXD

PERCENT DISTRIBUTION OF FAMILIES BY TYPE AND BY MONTHLY CHILD-CARE EXPENSES
1987 SIPP PANEL, WAVE 6

Type of Family and Monthly Child-Care Expenses	Reported Child-Care Expenses	Predicted Child-Care Expenses		Change in Percentage Points	
		Uncalibrated	Calibrated	Uncalibrated	Calibrated
Low-Income Families					
l-50	8.6	5.3	8.6	-3.3	**0.0**
51-100	15.5	21.6	26.4	6.1	10.9
101-150	16.5	24.4	22.2	7.9	5.7
151-200	21.5	19.6	20.6	-1.9	-0.9
201-250	13.2	11.8	7.7	-1.4	-5.5
251-300	6.5	3.4	3.7	-3.1	-2.8
301-350	8.6	4.2	3.4	-4.4	-5.2
351-400	2.9	2.5	2.1	-0.4	-0.8
401-450	1.9	1.9	1.7	0.0	-0.2
451+	5.0	5.2	3.7	0.2	-1.3
Total	100.0	100.0	100.0	0.0	0.0
Potentially Eligible Families					
l-50	16.9	7.8	11.0	-9.1	-5.9
51-100	8.9	29.1	33.8	20.2	24.9
101-150	23.6	26.8	23.9	3.2	0.3
151-200	18.2	14.3	17.0	-3.9	-1.2
201-250	13.3	11.5	7.5	-1.8	-5.8
251-300	7.1	3.5	1.1	-3.6	-6.0
301-350	4.0	1.6	1.4	-2.4	-2.6
351-400	1.4	1.4	1.7	0.0	0.3
401-450	0.0	1.7	0.0	1.7	0.0
451+	6.5	2.1	2.7	-4.4	-3.8
Total	100.0	100.0	100.0	0.0	0.0
Food Stamp Families					
l-50	17.2	**9.7**	17.8	-7.5	0.6
51-100	14.2	35.0	27.6	20.8	13.4
101-150	25.6	20.7	21.8	-4.9	-3.8
151-200	8.6	24.6	24.4	16.0	15.8
201-250	23.2	4.8	1.8	-18.4	-21.4
251-300	1.8	0.0	2.7	-1.8	0.9
301-350	0.0	3.5	1.4	3.5	1.4
351-400	4.0	0.0	1.2	-4.0	-2.8
401-450	0.0	0.0	0.0	0.0	0.0
451+	5.3	1.7	1.3	-3.6	-4.0
Total	100.0	100.0	100.0	0.0	0.0

SOURCE: 1987 Panel of SIPP, Wave 6.